Journal of
Clinical Psychology

OCTOBER 1958
Volume XIV Number 4

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Editorial Office: 5 Pearl Street
Brandon, Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
New York State Department of Mental Hygiene

MAURICE LORR
Veterans Administration Washington, D. C.

SAUL B. SELLS
School of Aviation Medicine Randolph Field, Texas

ROBERT I. WATSON
Northwestern University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street · · Brandon, Vermont

PSYCHOLOGICAL TEST SPECIALISTS

Exclusive publishers and distributors for selected psychological tests

TESTS NOW AVAILABLE

Kahn Test of Symbol Arrangement

Highly versatile projective test using symbolic plastic objects which are arranged and sorted. Validly discriminates among normals, neurotics, brain-injured, and psychotics, and indicates accompanying personality dynamics. **$25.00 for set of materials and manuals**

Full-Range Picture Vocabulary Test

Brief test of intelligence in two forms using the same 16 plates. Words are defined by choosing pictures. Valuable for intelligence screening, or with the physically handicapped. **$7.50 for set of materials**

Proverbs Test

Group and individual forms determine similarity of thought processes to those of various normal and clinical groups. Highly sensitive indicator of psychotic processes. **$4.50 for kit**

Vocational Apperception Test -- College Form

Separate male and female series of plates depicting college-related occupational activities, accompanied by brief explanation of use and interpretation as a projective test. **$7.25 for set of materials**

Famous Sayings

Agreement with proverbs, aphorisms and folk sayings is analyzed to determine personality structure. Of particular use in personnel selection. **$4.50 for kit**

Geist Picture Interest Inventory

Provides for expression of interests by choices among alternative pictured activities. Both qualitative and quantitative interpretations are possible. **$1.00 for sample set of materials**

Quick Picture Vocabulary Test

Each of three forms based on one test plate gives a good estimate of intellectual ability in 2-3 minutes. Can be used independently or as third form of FRPV. **$5.00 for set of materials**

TESTS TO BE PUBLISHED IN THE IMMEDIATE FUTURE

Emotional Projection Test

Emotional states are attributed to persons shown in 30 photographs. Extensive research has shown that responses to this test are extremely sensitive indicators of needs and feelings and can be effectively analyzed to determine personality structure.

Id-Ego-Superego Test

Subtests give independent estimates of personality components from a modified psychoanalytic point of view.

TESTS IN PREPARATION

Vocational Apperception Test-General Form; Electron Micrograph Apperception Test; It Scale for Children; Group Personality Projective Test; Projective Movement Sequences; Polychrome Index.

To order any of the tests now available, or to obtain information about those to be published, write:

PSYCHOLOGICAL TEST SPECIALISTS

Box 1441 MISSOULA, MONTANA

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. XIV | OCTOBER 1958 | No. 4

CONTENTS

STUDIES OF PERSONALITY DEVELOPMENT

Reproducible Personality Factors in Middle Childhood. RICHARD W. COAN and RAYMOND B. CATTELL - - - - - - - - - - - - - - - 339

Personality Factors in Nursery School Children as Derived from Parent Ratings. DONALD R. PETERSON and RAYMOND B. CATTELL - - - - 346

Personality Factors in the Readiness to Express Aggression. PHILIP WORCHEL - 355

The Relationship between the Self Concept and the Ideal-Self Concept as a Measure of Adjustment. IRWIN D. NAHINSKY - - - - - - - - - 360

The Self Concept in Paranoid Schizophrenia. ARTHUR H. ROGERS - - - 365

DIAGNOSTIC PROCESSES: VALIDATION STUDIES

The Pain Apperception Test: Psychological Correlates of Pain Perception. DONALD V. PETROVICH - - - - - - - - - - - - - - - - 367

Graphic Representation of Emotional Feelings. GEORGE A. PETERS and PHILIP R. MERRIFIELD - - - - - - - - - - - - - - - - - 375

The Validity of the Edwards Personal Preference Schedule in the Measurement of Dependency-Rebelliousness. MARVIN ZUCKERMAN - - - - - 379

Manifest Anxiety and Mother-Son Identification. LAWRENCE H. STEWART 382

The Psychasthenic and Hypomanic Scales of the MMPI and Uncertainty in Judgments. ALBERT V. GRIFFITH, HARRY S. UPSHAW and RAYMOND D. FOWLER - 385

Comparability of MMPI Card and Booklet Forms for Psychiatric Patients. ALBERT ROSEN, WILLIAM M. HALES and ROLAND M. PEEK - - - - - 387

A Note on the Effect on Self Description of Combining Anxiety and Hostility Items on a Single Scale. SAUL M. SIEGEL and MARVIN J. FELDMAN - - 389

Scores of Alcoholics on the Sixteen Personality Factor Questionnaire. NICHOLAS DE PALMA and HUGH D. CLAYTON - - - - - - - - - - - 390

The Leary Interpersonal Check List: A Reliability Study. RENATE G. ARMSTRONG - - - - - - - - - - - - - - - - - - - 393

DIAGNOSTIC PROCESSES: MEASURES OF INTELLECTUAL PERFORMANCE

Selective and Differentially Accelerated Intellectual Dysfunction in Specific Brain Damage. GRANVILLE C. FISHER - - - - - - - - - - 395

Intelligence Test Correlates of Shipley-Hartford Performance. LLOYD K. SINES - 399

A Proposed Method for Measuring Potential Intelligence. ALVIN R. MAHRER and LEWIS BERNSTEIN - - - - - - - - - - - - 404

CONTENTS—Continued

Inference of Intellectual Efficiency from the WAIS Vocabulary Subtest. STEPHEN L. FINK AND FRANKLIN C. SHONTZ - - - - - - - - - 409
Relationships between the Bender Gestalt Test and the Wechsler Intelligence Test for Children. ELIZABETH M. KOPPITZ - - - - - - - - 413
A Multiple Choice Supraordinality Test. HENRY N. PETERS - - - - - 416

PSYCHOTHERAPEUTIC PROCESSES: NATURE AND RESULTS

An Experimental Study of Transference and Resistance Phenomena in Psychoanalytically Oriented Psychotherapy. Moss L. RAWN - - - - 418
Insight and Prognosis. ALVIS ESKEY - - - - - - - - - - - - - 426
A Comparative Judgmental Study of "Defection" from Treatment at a Child Guidance Clinic. EUGENE E. LEVITT - - - - - - - - - - - 429

SELECTED CLINICAL STUDY

Alternating Two-Day Cyclic Behavior Changes. A. A. WEISS - - - - - 433

A Correction: A Comparison of Some Concepts of Psychological Health. BENJAMIN MEHLMAN and JANICE E. KAPLAN - - - - - - - - - 438

Editorial Opinion - - - - - - - - - - - - - - - - - - 439
Book Notices - 441
Subject Index, Vol. 14 - - - - - - - - - - - - - - - - - 443
Author Index, Vol. 14 - - - - - - - - - - - - - - - - - 448

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscriber's expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $8.50; Canadian subscriptions $9.00; foreign subscriptions $9.50 including postage; single copies $2.50. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post-office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont. Second-class postage paid at Rutland, Vt. and at additional mailing offices.

Copyright, 1958, Frederick C. Thorne, M. D.

REPRODUCIBLE PERSONALITY FACTORS IN MIDDLE CHILDHOOD[1]

RICHARD W. COAN AND RAYMOND B. CATTELL

University of Arizona *University of Illinois*

In a recent publication [3], the authors reported a factor-analytic study of the questionnaire responses of 151 first-and second-grade children. That research was undertaken with two major purposes in mind—the definition of personality structure at the six-to-eight-year level and the construction of a factored group questionnaire instrument which could be put to practical use. In order that any biases inherent in the reported factorization might be offset, a second independent analysis of questionaire data was undertaken.

Elsewhere we have reported findings derived from the analysis of rating and objective-test data obtained from a group of children in Decatur, Illinois [1, 2, 4]. This group, none of whose members were in our original questionnaire sample, also received the 200-item questionnaire which had been given to the earlier sample. The Decatur group consisted of all the pupils of four first-grade and four second-grade classes, a total sample of 198 children. This report deals with the responses of the 181 children for whom complete questionnaire data were available. These 181 children ranged in age from six years and three months to ten years, with a mean of seven years and five months and a standard deviation of 8.5 months. There were 90 boys and 91 girls. It is the purpose of the present article to report the analysis of the data for this sample and to indicate the reproducible factors that have emerged from a comparison with the results from our earlier study.

PROCEDURES

The items were identical with those described earlier [3]. They were divided into four blocks, each containing 50 items, and distributed over four sessions, generally on four consecutive school days. Each class was tested as a group. At each session, some objective tests were administered immediately after the questionnaire items. In all other respects, the testing procedure was like that described for the earlier sample. As before, the examiner read each item aloud, giving appropriate marking instructions, and the children indicated their responses on an answer form specially designed for the purpose.

The data were analyzed in much the same manner as were the earlier questionnaire data. An intercorrelation matrix of phi coefficients was obtained for the first 125 items. Fourteen items were then dropped by reason either of extremity of response cut or of an absence of significant correlations. These were not exactly the same items dropped in the original study, but there was considerable overlap in the selection of discarded items. The 111 surviving items were factored by the complete centroid method. A variety of criteria indicated about 18 significant factors, and the first 18 centroid factors were rotated to oblique simple structure. An initial oblimax solution was followed by graphic rotation.

Fifty marker items were selected to represent the 18 rotated factors. For each factor, there were between two and four markers which as a cluster effectively represented the given factor with irrelevant variance suppressed, or balanced out. The fifty markers were combined with items 126 to 200, and phi coefficients were again obtained. Again, 14 items were dropped, and the remaining 111 by 111 matrix was factored. Eighteen factors were retained for rotation. The markers served as reference variables for an initial rotation by the Thurstone analytic method. Subsequent graphic rotations carried the factors to a position consistent with a two-fold aim of attaining simple structure with replication of the factors obtained for the first half of the item battery.

[1]The research reported in this article was supported by a grant from the Department of Public Welfare of the State of Illinois.

The present rotations were independent of those done for the earlier sample. Within the present analysis, however, the second rotation was dependent on the first rotation. This procedure represents another application of the "matrix meshing" method earlier described [3]. The best indication of the successful alignment of the two present sets of factors is probably that shown by the loading pattern of each cluster of marker variables for the factor that it represents. A more objective assessment, however, is afforded by the total set of marker data. As in the earlier study, we intercorrelated factors from the two sets on the basis of the marker data. For this purpose, we treated each marker variable as a person and the corresponding set of reference-vector correlations as scores. The results are shown in Table 1. Even by

TABLE 1. CORRELATIONS BETWEEN FACTORS FROM TWO HALVES OF BATTERY

Factor in first analysis	Factor in second analysis																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	61	16	-05	-71	-15	06	13	25	-16	-03	-41	-22	22	31	-24	-24	-16	-01
2	-12	69	-29	-02	-13	-04	03	24	-33	-06	-09	11	-06	-13	-13	08	-16	15
3	05	-07	80	15	27	30	06	-11	13	31	-14	02	02	-07	03	-10	35	-20
4	-30	10	-01	84	12	18	07	-01	10	-18	16	40	08	-18	15	07	07	-02
5	-50	-11	-07	23	57	08	-03	-06	16	-04	12	16	30	-25	15	-17	09	-03
6	-04	-19	17	-11	35	65	-02	-21	-05	18	03	37	16	22	29	15	20	17
7	14	-21	09	12	-03	01	86	-18	-32	06	-04	-32	05	06	31	-03	13	13
8	03	05	-41	03	07	-00	23	62	12	19	02	-17	12	-29	-06	07	-13	-01
9	01	15	-04	-07	05	-28	-03	16	67	04	20	04	22	-05	09	01	-13	01
10	40	17	-11	-39	-25	05	-05	13	-01	58	04	-10	-23	06	-36	-03	-10	16
11	-17	-43	-03	-08	41	07	16	-31	-34	12	68	02	-19	24	-24	35	-10	17
12	19	-16	-14	15	02	14	-19	-16	-08	-07	36	64	-05	-10	10	37	-11	42
13	-16	42	-03	12	-34	09	-01	-20	29	-00	-12	-08	68	30	20	21	22	01
14	36	12	-04	-11	-23	05	02	-32	-08	-28	25	-11	-16	53	-05	-12	-30	09
15	-32	-16	17	-18	12	-06	17	-26	-08	11	05	11	11	-37	37	33	-03	-29
16	-14	-20	-08	24	-07	06	-26	-23	00	-09	41	20	08	-05	22	78	28	-26
17	04	-24	08	10	05	03	11	-10	-27	08	-06	12	26	04	21	19	74	22
18	11	-15	-20	00	-00	01	03	-26	15	09	22	23	17	32	36	-08	16	57

this somewhat defective criterion, it can be seen that the factors from the two halves of our data match with little ambiguity. Corresponding factor numbers in the table, of course, represent corresponding factors. For present consideration, the rotated factors have been arranged in order of total variance. Magnitude of variance was assessed on the basis of the two rotated matrices combined.

FINDINGS

A complete exposition here of the factor data from this study would demand more space than is warranted by our immediate concerns.[2] We shall instead focus our attention on areas of agreement between present findings and those of our earlier study, and we shall offer preliminary interpretations for the reproducible factors that have emerged. In a later publication [5], we shall have an opportunity to examine additional data relevant to the psychological nature of the questionnaire factors derived from the two studies.

One way of making an objective comparison of the factors in the two studies is to correlate them on the basis of loadings or correlations for the common variables. A procedure like that used in comparing the halves was employed. Each of the 165 variables common to the two studies was treated as a person with scores (reference-vector correlations) on a series of factors. Within a study the two sub-columns for each factor (derived from the two halves of the data) were treated as a single column. The two values for each marker variable were averaged, and their mean was accepted

[2]The centroid matrices, transformation matrices, and rotated reference-vector matrices for this study have been deposited with the American Documentation Institute. Order Document No. 5706 from American Documentation Institute, 1719 N Street N. W., Washington 6, D. C., remitting $1.75 for 35 mm. microfilm or $2.50 for 6 by 8 in. photocopies.

as the score. On this basis, the 18 factors from our earlier study were correlated with those in the present study. The results are shown in Table 2.

TABLE 2. CORRELATIONS BETWEEN PRESENT FACTORS AND THOSE FROM EARLIER STUDY

Earlier factor	Present factor 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	-44	-13	20	46	04	12	-07	19	05	-07	05	28	01	-05	-05	20	03	-05
2	-34	-26	41	21	12	18	-01	-24	-07	-27	06	22	08	01	29	06	11	-01
3	-00	09	35	12	04	11	-01	-11	-02	-06	-13	17	-02	-06	17	-02	-06	-12
4	05	35	-11	-04	-01	-11	-07	-15	20	05	-05	-16	06	06	-17	-11	-20	-05
5	-05	-27	15	07	05	25	16	-17	-11	02	20	12	13	26	18	19	08	-05
6	17	-10	-16	-08	15	07	04	29	-13	13	-01	-03	-14	-07	-07	-15	14	12
7	14	07	-32	-07	09	-03	03	14	01	17	-02	20	-19	-11	-00	-04	01	02
8	-10	-02	02	-10	-00	14	03	04	07	-07	-14	-06	01	-18	-01	05	04	-00
9	-11	07	-09	19	05	-07	01	07	-05	-03	-01	-13	05	04	03	-17	02	10
10	33	05	-10	-28	02	04	23	01	-15	22	-10	-05	04	12	08	01	03	-01
11	-26	-15	08	27	04	-04	-10	-21	-17	-24	26	04	10	15	-12	25	-02	01
12	04	10	-15	04	-08	-18	-09	-10	04	-20	-15	02	08	08	-18	-05	-03	-03
13	-20	-03	05	10	00	-01	03	05	-11	-09	09	17	-06	07	05	09	04	03
14	-04	-06	05	13	14	07	-01	-19	-02	06	19	11	05	05	22	-01	-14	-13
15	14	-03	03	03	-02	-16	26	-12	-06	05	-12	-18	-05	-07	-06	-20	-06	05
16	08	-10	06	-12	04	13	10	06	-17	-01	03	08	09	-02	06	11	09	-03
17	08	28	-11	-13	-10	-18	00	-11	03	09	-09	-10	06	-08	-16	-03	-09	07
18	03	11	-07	01	13	-00	06	04	-03	-01	-11	-12	-03	-19	01	-05	-08	03

For a single coefficient, the requisite levels for significance at the 5% and 1% levels would be roughly .15 and .20. It is evident from Table 2 that there is considerable correspondence between the factors of the two studies. In particular, it may be noted that each of the first four factors in either study relates significantly to several factors in the other study. In this general sense, there is some replication of the original factors. If we had only the correlations as a guide, however, we should have some difficulty in aligning the two sets of factors.

The correlation coefficient, unfortunately, is not an ideal index of factor equivalence. With a well rotated factor, it gives undue emphasis to the minor variations among the great bulk of the variables lying in the hyperplane. Where factors vary in overall magnitude of variance, it also obscures the order of relationships, favoring large-variance factors. Thus, we know from other evidence that present factor 15 aligns best with earlier factor 14, though it correlates more highly with factor 2 from the earlier data. The highest correlations are among the factors at the beginning of the two sets, since the factors from each study are arranged in order of magnitude. Despite this fact, we can see some overall tendency for factors in either set to correlate most highly with factors of similar position (relative magnitude) in the other set.

In undertaking this study, however, we were less interested in simply demonstrating replicatability than in identifying replicatable factors. We had the additional practical goal of constructing a factored questionnaire which could rest soundly on two independent assessments of factor structure. The extent to which any given factor in one study appears to relate to a factor in the other study depends on the index one chooses. The more important question is that of the factor which best aligns with a given factor across the two samples. Since there is no single criterion which will provide a reliable answer in every case, some subjective evaluation is inevitable. To insure a reasonable degree of objectivity in our matching of factors, we compared findings based on each of the four following criteria:

a. The correlations shown in Table 2,

b. The similarity of the profile (column or row) of correlations in a complete 36 by 36 intercorrelation matrix for the factors from the two studies (an indication of proximity in factor space which is more independent of the length of the vector),

c. Similar magnitude with respect to total variance accounted for by the factor, and

d. The possession of similarly loaded salient variables.

This combination of criteria afforded virtually unambiguous matching for eleven factors. The matching factor pairs were as follows:

Present factor	Earlier factor
1	2 (–)
2	4
3	3
4	1
6	5
7	15
8	6
9	10 (–)
10	11 (–)
13	7 (–)
15	14

These factor pairs have been incorporated into a factored questionnaire. In a later publication [5], we shall describe the construction of that questionnaire in more detail and present evidence bearing on the nature of the component factor measures. None of the seven remaining factors in either study clearly corresponds to a factor in the other study. This discrepancy may be partly attributable to sampling error. It is also possible that the few variables retained for analysis in either study but not in the other had some effect on the factor resolution. No doubt, it would have been possible to secure much better agreement between the studies with a tolerable loss of simple structure if the procedure of blind independent rotation had not been followed. This procedure, however, enables us to concentrate our attention on the more stable factors which can be expected to reappear in subsequent research.

To give the reader some preliminary insights into the nature of the stable factors, we may examine some of the salient items for each one. It is not contended, of course, that the content of the items is alone sufficient for interpretation, though this content usually shows an obvious consistency with interpretations based on external evidence. In later publications, we shall examine correlations between our questionnaire factors and variables derived from other instruments. For the present, we shall confine our attention to the content of highly loaded salient items. We shall include an item in each factor list only if it has a loading (strictly speaking, a reference-vector correlation) of at least .30 in one study accompanied by a loading of at least .20 in the other study.

First reproduced factor (present factor 1, earlier factor 2–)

3. Would you rather play: (a) school, or (b) *cowboys and Indians?* (present loadings: .56, .41; earlier: –.06, –.30)[3]

47. Which would you rather be: (a) a teacher, or (b) *a doctor?* (present: .56; earlier: –.35, –.07)

75. Would you rather: (a) be a teacher, or (b) *fly an airplane?* (present: .68, .32; earlier: –.28)

92. Do you like to talk to teachers? (a) yes, or (b) *no.* (present: .32; earlier: –.31)

110. If you were in a play, would you rather be: (a) *a hunter,* or (b) a teacher? (present: –.57, –.31; earlier: .30)

113. Would you rather: (a) be in a play, or (b) *make something out of wood?* (present: .44; earlier: –.30)

147. Do people ever call you naughty and mischievous? (a) *yes,* or (b) no. (present: –.31; earlier: .24)

The negative end of this factor seems to be characterized by a form of socialization manifested especially in an acceptance of the school situation. At the positive end, there is a greater preference for non-social or non-academic pursuits. Both item content and correlations we have found with other measures point to an identifica-

[3]The answer corresponding to the positive pole of each factor is italicized. For each item that served as a marker variable in either study, the loadings in both factorizations within the study are given.

tion of this factor with either adult factor A- (Schizothymia vs. Cyclothymia) or adult factor K- (Dislike of Education vs. Socialized Morale).

Second reproduced factor (present factor 2, earlier factor 4)

25. Do other children ever pick on you? (a) *yes*, or (b) no. (present: –.36; earlier: –.05, –.20)

39. (a) *Do you ever get sad*, or (b) are you happy all the time? (present: –.42, –.53; earlier: –.26)

103. Do you ever feel like crying when you see something sad in the movies? (a) *yes*, or (b) no. (present: –.23, –.23; earlier: –.09, –.32)

142. Does it bother you if people say you don't listen? (a) *yes*, or (b) no. (present: –.25; earlier: –.32)

147. Do people ever call you naughty and mischievous? (a) *yes*, or (b) no. (present: –.32; earlier: –.32)

157. Do people sometimes punish you when you haven't done anything wrong? (a) *yes*, or (b) no. (present: –.39; earlier: –.24)

189. (a) *Do your friends sometimes say bad things about you*, or (b) do they only say true things? (present: –.53; earlier: –.31)

190. Do you ever do things that you should not do? (a) *yes*, or (b) no. (present: –.43; earlier: –.21)

192. Do you sometimes feel a little scared when you're up on a high place? (a) *yes*, or (b) no. (present: –.31; earlier: –.21)

198. Do grown-ups ever say you daydream too much? (a) *yes*, or (b) no. (present: –.41; earlier: –.33)

This factor seems to present a consistent picture of general emotional distress or discomfort. In all likelihood, it corresponds to known factor O (Anxious Depression), which at this age level may be inseparable from factor Q_4 (Nervous Tension).

Third reproduced factor (present factor 3, earlier factor 3)

26. Would you rather: (a) *go to school*, or (b) go on a long trip in the car? (present: –.58, –.55; earlier: –.47, –.56)

29. Do you think: (a) *school is easy*, or (b) school is hard? (present: –.26, –.19; earlier: –.34)

94. Would you rather: (a) *go to school*, or (b) work at home? (present: –.49, –.45; earlier: –.50, –.49)

98. Do you like school: (a) *a lot*, or (b) not very much? (present: –.46; earlier: –.35)

135. Would you rather: (a) *learn something in school*, or (b) watch a game? (present: –.21; earlier: –.45)

183. Which do you like better: (a) *the books you have in school*, or (b) funny books? (present: –.40; earlier: –.28)

In its stress on an acceptance of school, this factor seems related to the first factor, perhaps reflecting a different aspect of socialization. There is possibly a bit more in the item content of this factor that suggests an incorporation of values associated with school, rather than simply a liking for school and teacher. Other evidence supports an identification of this factor as factor G (Superego Strength).

Fourth reproduced factor (present factor 4, earlier factor 1)

42. Do you like to climb trees? (a) yes, or (b) *no*. (present: .48, .51; earlier: .59, .52)

70. Would you rather: (a) climb a tree, or (b) *look at a book?* (present: .45, .60; earlier: .48)

75. Would you rather: (a) *be a teacher*, or (b) fly an airplane? (present: –.19, –.30; earlier: –.30)

110. If you were in a play, would you rather be: (a) a hunter, or (b) *a teacher?* (present: .15, .36; earlier: .32)

116. Would you rather: (a) climb a tree, or (b) *color a book?* (present: .54, .51; earlier: .42, .47)

136. Would you rather: (a) hunt for birds, or (b) *draw pictures of birds?* (present: .32; earlier: .45)

This factor shows some resemblance to the first and third factors. The common features, however, are clearly secondary to a sort of timidity which makes quiet,

culturally sanctioned pursuits preferable to more adventurous ones. This factor corresponds to adult factor I (Premsia or Emotional Sensitivity).

Fifth reproduced factor (present factor 6, earlier factor 5)

10. (a) *Do you like for other children to play with your toys,* or (b) does it sometimes bother you? (present: –.32; earlier: –.29)

48. Would you rather be: (a) *a grown-up,* or (b) a baby? (present: –.30; earlier: –.07, –.26)

114. Do you cry: (a) more than other boys and girls, or (b) *less?* (present: .40, .11; earlier: .15, .21)

168. Do people boss you around too much? (a) yes, or (b) *no.* (present: .30; earlier: .28)

This appears to be a reflected version of factor D (Infantile Emotionality vs. Phlegmatic Frustration Tolerance). The high scorer professes an acceptance of reasonable restrictions, while the low scorer appears to lack satisfactory emotional control.

Sixth reproduced factor (present factor 7, earlier factor 15)

1. Which do you like better: (a) cats, or (b) *dogs?* (present: .61, .60; earlier: .39, .51)

49. Which would you rather be: (a) *a dog,* or (b) a cat? (present: –.52, –.52; earlier: –.49, –.47)

Except for the "dog-and-cat" items, this factor shows a conspicuous lack of highly loaded items. There is no clear correspondence to any known factor. Inspection of the many items with more moderate loadings, however, suggests an interpretation of "motor expressiveness" vs. "socialized restraint."

Seventh reproduced factor (present factor 8, earlier factor 6)

35. Who usually has better ideas: (a) *you,* or (b) your friends? (present: –.32; earlier: –.30)

55. Can you do things: (a) *better than most boys and girls,* or (b) not as well as most boys and girls? (present: –.32, –.18; earlier: –.21)

69. Would you rather: (a) *look at a picture book by yourself,* or (b) look at it with another boy or girl? (present: –.32; earlier: –.26)

84. (a) *Do you always do what you want to,* or (b) do you have to do things you don't want to do? (present: –.32; earlier: –.28, –.36)

93. Would you rather: (a) *build things by yourself,* or (b) build things with your friends? (present: –.44, –.48; earlier: –.26)

111. Do you like to: (a) *tell other children what to do,* or (b) do what other children want to do? (present: –.47, –.29; earlier: –.21)

125. (a) *Are you stronger than other children,* or (b) are they stronger than you? (present: –.27; earlier: –.31, –.06)

155. Would you rather: (a) *make something with blocks or bricks,* or (b) play with other children? (present: –.38; earlier: –.21)

This clearly corresponds to adult factor E (Ascendance vs. Submissiveness). It is typical of factor E as it appears in rating and questionnaire studies at lower age levels that the emphasis in content is shifted from dominance behavior to aggressive, independent self-assertion. The tendency of the high scorer to prefer solitary activities to cooperative social activities is consistent with this picture.

Eighth reproduced factor (present factor 9, earlier factor 10–)

37. Do you like to: (a) run, or (b) *just walk?* (present: .24; earlier: –.31)

The items with lower loadings seem, with item 37, to conform to a pattern of quietness, combined perhaps with seclusiveness and sullenness. The factor may be related to A (Cyclothymia vs. Schizothymia), but further information seems to be needed for adequate identification.

Ninth reproduced factor (present factor 10, earlier factor 11–)

24. Would you rather talk to: (a) *your father,* or (b) your mother? (present: –.32; earlier: .24, .14)

27. Which would you like better: (a) to hear stories about bears, or (b) *to have bears here right now?* (present: .20; earlier: –.24, –.30)

80. When you want to say something: (a) do grown-ups always listen to you, or (b) *do they do all the talking?* (present: .26; earlier: –.34)

89. Would you rather: (a) *listen to the radio and TV*, or (b) go to a party? (present: –.32, –.20; earlier: .09, .25)

This looks vaguely like a reflected version of factor I (Premsia or Emotional Sensitivity), but it cannot really compete with the fourth factor as a match for factor I. No other interpretation seems clearly supported.

Tenth reproduced factor (present factor 13, earlier factor 7–)

55. Can you do things: (a) better than most boys and girls, or (b) *not as well as most boys and girls?* (present: .27, .04; earlier: –.32)

139. Does your mother ever go away for a little while and leave you at home all by yourself? (a) yes, or (b) *no.* (present: .26; earlier: –.34)

158. Do you ever talk in your sleep and wake yourself up? (a) yes, or (b) *no.* (present: .29; earlier: –.31)

This factor, too, is not readily interpretable on the basis of present evidence.

Eleventh reproduced factor (present factor 15, earlier factor 14)

9. Do you like: (a) *to meet new boys and girls*, or (b) to be with children you already know? (present: –.27; earlier: –.46, –.31)

23. When children fight, do you think grown-ups sometimes punish the wrong one? (a) *yes*, or (b) no. (present: –.21; earlier: –.30, –.08)

69. Would you rather: (a) look at a picture book by yourself, or (b) *look at it with another boy or girl?* (present: .25; earlier: .33)

77. Would you rather: (a) look at funny books, or (b) *talk to a friend?* (present: .21; earlier: .34, .50)

The overall trend of friendly sociability would seem to make this the best match within the present set for factor H (Adventurous Cyclothymia vs. Withdrawn Schizothymia).

Summary

A 200-item questionnaire was administered to 181 first- and second-grade children. Two overlapping factor analyses, utilizing responses to 172 of the items, yielded 18 obliquely rotated factors (or aligned factor pairs). These factors were compared with 18 factors obtained in a prior independent study. Eleven factors from either study were found to correspond reasonably well with 11 factors in the other study. Examination of the apparently stable factors reveals trends related to each of the following factors known from studies of older subjects: A, D, E, G, H, I, K, O, and Q_4. Later publications will report the construction of a factored questionnaire instrument based on our present work and relations found between our stable factors and factors derived from other kinds of data.

References

1. Cattell, R. B. and Coan, R. W. Child personality structure as revealed in teachers' behavior ratings. *J. clin. Psychol.*, 1957, *13*, 315-327.
2. Cattell, R. B. and Coan, R. W. Objective-test assessment of the primary personality dimensions in middle childhood. (In Preparation)
3. Cattell, R. B. and Coan, R. W. Personality dimensions in the questionnaire responses of six and seven year olds. *Brit. J. educ. Psychol.* (In Press)
4. Cattell, R. B. and Coan, R. W. Personality factors in middle childhood as revealed in parents' ratings. *Child Developm.*, 1957, *28*, 439-458.
5. Coan, R. W. and Cattell, R. B. The development of the Early School Personality Questionnaire. (In Preparation)

PERSONALITY FACTORS IN NURSERY SCHOOL CHILDREN AS DERIVED FROM PARENT RATINGS[1]

DONALD R. PETERSON AND RAYMOND B. CATTELL

University of Illinois

A great deal has been written, but little is known about the development of personality in early childhood. One impediment to the accumulation of such knowledge has been a basic lack of adequate measurement techniques, and a still more fundamental vagueness in structural definition of major personality dimensions in subjects at various age levels. The present study is part of a much larger project designed to obtain descriptive information on personality structure among children in three different age ranges, and to relate that information to previously established knowledge about personality structure in adults. Results of studies on children between the ages of 10 and 14[5, 6, 7] and of six- and seven-year-old children[3, 4] have already been reported.

This investigation is concerned with children four and five years of age. In general design, it is similar to the previous studies, but the scope of the data gathering operation has been somewhat expanded. As before, objective tests have been administered and responses to questionnaire items elicited, but instead of relying on just one or two sets of ratings, three kinds of "life record" data have been gathered, *viz.*, ratings by teachers, ratings by parents, and time sampling observations. Separate factorizations have been carried out and the results of each analysis integrated with those previously found in older subjects. Results for teacher ratings[10] and time sampling observations[9] have already been reported. The present article concerns ratings by parents, which are less cleanly objective than actual behavior observations, or even teacher ratings, but which are based on a long and intimate relationship which no other observers can duplicate.

Subjects and Procedures

Over a period of a year, data were gathered on 32 boys and 48 girls between the ages of 4-0 and 6-0 (X = 4-7). All were pupils at the University of Chicago Nursery School. They were considerably above the national and community averages in respect to both intelligence and socio-economic status. Sampling deviations of this kind, however, are irrelevant to factor generality so long as representative variance along major personality dimensions is maintained, and the results of previously reported analyses[8, 9, 10] suggest that this second more pertinent sampling criterion has been adequately met.

A 43-item rating form was constructed with a view toward comprehension of as broad a range of personality variables as parents might be expected to judge with reasonable precision. In selecting variables, we desired especially to include a number of activities which parents alone are competent to assess, *e.g.*, behavior at bedtime, feeding problems, fear of the dark. The inclusion of these, however, required omission of others. Practical limitations on length necessitated exclusion of many familiar marker variables, and while this exposed the study to difficulties in factor interpretation which will soon be apparent, it seemed strategically wise to concentrate on such family-relevant variables, and to depend on later analyses and subsequent investigation for more definitive interpretation than the present mode of variable selection permits. The 43 variables are presented, in condensed statement, below, along with a 44th variable based on teacher report which was eventually included in the analysis.[2]

---

[1]This project was supported by the University Research Board, University of Illinois. We are grateful to the staff of the University of Chicago Nursery School and to the cooperating parents for their help in gathering the data, and to Peter V. Vygantas and Donald Tollefson for their help in the data analysis.

[2]The complete rating schedule can be obtained by writing to the authors.

1.	Prefers younger playmates	*vs.*	Prefers older playmates
2.	Not destructive	*vs.*	Destructive
3.	Selfish	*vs.*	Generous
4.	Fearful	*vs.*	Fearless
5.	Taciturn	*vs.*	Talkative
6.	Active	*vs.*	Inactive, Bored
7.	Inquisitive about sex	*vs.*	Not inquisitive about sex
8.	Seldom cries	*vs.*	Often cries
9.	Seldom complains	*vs.*	Often complains
10.	Dislikes school	*vs.*	Likes school
11.	Quiet in speech	*vs.*	Loud in speech
12.	Often uses forbidden words	*vs.*	Seldom uses forbidden words
13.	Seldom sulks or pouts	*vs.*	Often sulks or pouts
14.	Has many accidents	*vs.*	Has few accidents
15.	Has little physical endurance	*vs.*	Has great physical endurance
16.	Speaks fluently	*vs.*	Stutters
17.	Articulates well	*vs.*	Articulates poorly
18.	Resists going to bed	*vs.*	Eager to go to bed
19.	Under-achieves in school	*vs.*	Over-achieves in school
20.	Never runs away from home	*vs.*	Often runs away from home
21.	Often tells lies	*vs.*	Never tells lies
22.	Never steals	*vs.*	Steals
23.	Afraid of animals	*vs.*	Unafraid of animals
24.	Unafraid of strangers	*vs.*	Afraid of strangers
25.	Afraid of the dark	*vs.*	Unafraid of the dark
26.	Never bites fingernails	*vs.*	Bites fingernails
27.	Sucks thumb	*vs.*	Never sucks thumb
28.	Never masturbates	*vs.*	Masturbates
29.	Has frequent stomach upsets	*vs.*	Seldom has stomach upsets
30.	Has frequent eliminative disturbances	*vs.*	Seldom has eliminative disturbances
31.	Seldom has colds	*vs.*	Frequently has colds
32.	Seldom has headaches	*vs.*	Frequently has headaches
33.	Has frequent temper tantrums	*vs.*	Seldom has temper tantrums
34.	Talked early	*vs.*	Talked late
35.	Domineering	*vs.*	Not domineering
36.	Toilet trained early	*vs.*	Toilet trained late
37.	Solemn	*vs.*	Cheerful
38.	"Shows off"	*vs.*	Seldom "shows off"
39.	Compliant	*vs.*	Not compliant
40.	Seldom daydreams	*vs.*	Daydreams excessively
41.	Unhappy unless with others	*vs.*	Prefers to be alone
42.	Shows little affection	*vs.*	Affectionate
43.	Refuses to eat certain foods	*vs.*	Eats anything served
44.	Not enuretic	*vs.*	Enuretic

Judgments were originally made on a three-point scale. These were dichotomized by combining the two categories of lowest frequency, phi coefficients were computed, and 14 factors extracted by the complete centroid method. An over-all analytic rotation was performed by means of the quartimax routine for the electronic computer, and a lengthy series of additional oblique rotations was then conducted with the aid of graphic plots. Near the end of the latter operation, salient variables were inspected and several further rotations performed in the light of knowledge of variable content. At the same time, blind rotations were pursued to a definitive solution. The factors presented below represent, in our judgment, the best single solution to emerge from both kinds of rotational procedure.[3] In the presentation, we shall list variables with loadings of .20 and above for each factor, and in accordance with convention reverse each bipolar designation as required by sign so that the phrase or term given first for each item will correspond with the positive pole of the factor.

The degree of resemblance between the present factors and those found in the only other analysis of parent ratings in the entire project, *i.e.*, the study by Cattell and Coan of six- and seven-year-olds [4] has been estimated by intercorrelating load-

[3]Hyperplanes for all but three of the factors (Factors 3, 6, and 13 as presented below) are significant at the .05 level by Bargmann's test, and hyperplanes for those factors closely approach significance.

ings on 38 variables examined in both studies[2]. The correlations which result from application of this procedure measure resemblance between loading pattern for each factor in the present study and loading patterns for similarly identified factors in the previous study of older children.

Results

Loadings on primary reference factors are presented in Table 1. Correlations between the vectors are given in Table 2.[4] Correlations with loadings obtained in the study of parent ratings on six- and seven-year-olds[4] will be reported in the discussion of each factor.

Table 1. Rotated Reference Vector Matrix

Variable	Factor Number 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.	07	-01	01	-11	27	38	-09	-05	-01	-35	17	08	06	01
2.	30	-02	-04	-01	05	24	10	-04	11	-22	-06	-36	06	01
3.	07	02	-11	-10	-46	-08	01	-17	11	-29	07	02	-18	-30
4.	-04	08	27	04	08	13	36	-10	21	-08	02	-02	10	-25
5.	04	-19	-01	02	-34	-14	27	27	-23	06	-01	-02	23	-01
6.	13	02	10	03	30	-06	-21	-22	02	36	-01	20	05	16
7.	-33	-08	06	-05	01	25	03	01	05	-04	-45	-04	-06	-02
8.	52	-07	-16	-05	-03	-09	-10	-00	-01	-07	24	01	-11	33
9.	02	-24	-19	17	04	-13	04	-04	03	25	14	-05	-13	46
10.	-09	-17	11	00	-10	09	-07	-15	-20	-02	07	-58	11	-03
11.	07	-01	-27	07	02	19	04	29	-00	07	07	-23	45	-03
12.	-30	18	05	-07	-19	10	30	-05	05	01	-09	09	01	-20
13.	25	-03	-12	04	01	09	07	-17	-20	-23	-41	04	11	03
14.	-55	06	-02	41	-05	11	-06	-11	09	-12	-04	11	-04	-06
15.	-03	-06	02	06	08	28	-07	03	02	-53	-05	02	-26	09
16.	-06	68	-02	04	-06	05	-05	-06	06	-08	-20	04	08	-23
17.	-01	66	10	-02	-03	06	01	05	09	12	-17	15	10	-22
18.	01	-03	04	-29	-04	13	-07	-65	-01	08	-06	-31	-22	-04
19.	-03	-09	-01	06	03	06	-00	-06	09	-06	05	-03	01	-25
20.	-03	02	11	61	07	-20	-04	08	03	-08	-04	06	01	15
21.	-62	-09	-07	-04	01	06	-04	-01	05	-02	01	11	-12	-01
22.	12	04	-07	03	06	-01	-03	-01	00	-02	10	-52	20	09
23.	-04	09	31	13	-26	-23	18	01	00	-17	24	07	-28	-01
24.	-06	28	-06	-21	46	42	-02	09	11	01	04	16	-05	09
25.	-06	-04	61	07	-03	-01	-04	-03	-13	-02	13	-08	-03	05
26.	-03	04	-20	03	-28	-18	16	-03	-24	-03	20	-24	24	-26
27.	-07	00	-01	-05	-06	16	-02	-06	-28	-07	-06	-09	30	00
28.	33	-01	-08	-07	-02	09	-41	-08	07	04	06	-06	-10	02
29.	08	-05	55	-06	00	07	-18	05	13	05	-12	-07	-02	-10
30.	01	-05	52	08	14	-04	07	-01	-08	05	-09	-05	-10	01
31.	07	-02	-01	-31	-08	-02	56	05	10	-04	03	90	-10	-18
32.	03	-11	-11	-18	05	-13	26	28	05	08	23	-12	14	-19
33.	-12	-04	-05	-20	-05	09	07	01	-03	-18	54	-10	-24	-01
34.	-06	-51	09	-08	-13	-24	-05	25	13	02	-04	07	-02	-02
35.	-08	07	-02	09	-01	-01	11	-60	09	-00	02	-02	-17	-01
36.	03	-01	-09	-10	-00	47	-03	12	30	06	04	-15	-17	09
37.	09	00	18	-10	-25	-04	13	-09	-01	-44	05	-07	-13	-19
38.	-19	-33	-11	07	-04	07	-03	-18	70	03	-01	-05	-36	05
39.	29	-29	-11	04	52	10	03	-05	-03	-08	01	02	-04	29
40.	09	-06	-01	68	-03	20	-00	-06	03	06	-06	-08	26	-02
41.	08	03	11	01	01	13	02	11	51	-02	09	13	05	-28
42.	05	02	03	08	-27	-21	-24	05	-01	-34	07	03	01	-06
43.	-10	09	22	-07	00	07	-41	09	03	-02	-30	06	04	03
44.	-00	-06	09	-01	02	33	10	01	00	02	-22	08	13	11

[4]The correlation matrix, unrotated factor matrix, quartimaxed factor matrix, and transformation matrix have been deposited with the American Documentation Institute. Order Document No. 5702 from American Documentation Institute, 1719 N St., NW, Washington 6, D. C., remitting $1.25 for 35 mm. microfilm or $1.25 for 6 x 8 inch photocopies.

TABLE 2. INTERCORRELATIONS AMONG REFERENCE VECTORS

Factor Number	Factor Number													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.		06	08	-10	-05	-18	-11	-28	-03	-20	29	-03	-02	-12
2.			11	-09	-19	-04	05	-07	-18	-07	05	21	19	-49
3.				-01	10	-01	10	04	-05	11	-23	06	-01	-15
4.					05	-23	-09	03	-03	05	-16	06	31	19
5.						42	02	05	04	16	-03	22	-00	48
6.							10	00	-16	-15	-14	-12	09	11
7.								12	-07	-03	05	00	12	-25
8.									09	35	-02	22	22	17
9.										02	-01	16	-29	-21
10.											-17	-01	35	15
11.												-13	-18	01
12.													-20	06
13.														-25

FACTOR 1. SUPER EGO STRENGTH (G)

Item No.	Loading	*Description of Variable*	
21	-62	Seldom tells lies	*vs.* Often tells lies
14	-55	Seldom has accidents resulting in injury	*vs.* Often has accidents resulting in injury
8	52	Seldom cries	*vs.* Often cries
7	-33	Not inquisitive about sex	*vs.* Inquisitive about sex
28	33	Never masturbates	*vs.* Masturbates
2	30	Not destructive	*vs.* Destructive
12	-30	Seldom uses forbidden words	*vs.* Often uses forbidden words
39	29	Compliant	*vs.* Not compliant
13	25	Seldom sulks or pouts	*vs.* Often sulks or pouts

The consistent social desirability of all salient variables supports identification of this first and largest factor as G. Correlation with loadings on G as isolated in the study of six-and seven-year-olds, however, is only .29. The most obvious discrepancy between the present factor and its appearance at older age levels, *viz.*, the absence here of perseverance, determination, and responsibility, is most likely a direct consequence of failure to include any of these variables in the analysis. Whether responsibility and perseverance would reappear with characteristically heavy saturations if they *were* examined in conjunction with the variables actually studied is a matter of speculation, or better, empirical test, but the likelihood of such an occurrence is enhanced by the results on seven-year-olds, where perseverance and responsibility were included along with the present variables, and a considerably more typical Super Ego pattern emerged.

FACTOR 2. COMENTION VS. ABCULTION (K)

16	68	Speaks fluently	*vs.* Stutters
17	66	Articulates well	*vs.* Articulates poorly
34	-51	Learned to talk early	*vs.* Learned to talk late
38	-33	Seldom shows off	*vs.* Frequently shows off
24	28	Unafraid of strangers	*vs.* Afraid of strangers
39	-29	Not compliant	*vs.* Compliant
9	-24	Often complains	*vs.* Seldom complains

A correlation of .53 between the loadings on Factor 2 and those for K as previously identified with seven-year-olds[4] bespeaks reasonably close agreement. The strong verbal component which marks the factor both for these and for older children suggests a developmental relationship between linguistic skill and the readiness to accept cultural influence which probably constitutes the essential nature of K. Disobedience and proneness to complain, which at first seem inconsistent with this interpretation of K, probably represent the independent-mindedness which has been found quite regularly in the adult pattern, and which serves in the occasionally difficult differentiation of K from B and G. In the latter two factors, disobedience and readiness to complain should load negatively or not at all.

FACTOR 3. GUILT PRONENESS VS. CONFIDENCE (O)

25	61	Afraid of the dark	*vs.* Unafraid of the dark
29	55	Has frequent stomach upsets	*vs.* Seldom has stomach upsets
30	52	Has frequent eliminative disturbances	*vs.* Seldom has eliminative disturbances
23	31	Afraid of animals	*vs.* Unafraid of animals
11	-27	Loud in speech	*vs.* Quiet in speech
4	27	Fearful	*vs.* Fearless
43	22	Refuses to eat certain foods	*vs.* Eats anything served

A modest but significant correlation (r = .38) was found between the present loadings and those established for O in seven-year-olds. As in the older children, fearfulness of a very general kind is predominant, but the depressive component, which appeared only as a readiness to cry in seven-year-olds, is not evident at all in these younger children. The subjective sense of inadequacy prominent in the adult pattern was found in neither group of children, and developmental progression from relatively pure fearfulness (including its visceral expression), followed by a conjunction of fearfulness and depression, and eventuating in the more complex adult pattern of fearfulness, depression, and feelings of inadequacy, is suggested as an hypothetical guide for further empirical study.

FACTOR 4. EXCITABILITY-INSECURITY (D)

40	68	Seldom daydreams	*vs.* Daydreams excessively
20	61	Seldom runs away	*vs.* Often runs away
14	41	Has many accidents resulting in injury	*vs.* Seldom has accidents resulting in injury
31	-31	Frequently has colds	*vs.* Seldom has colds
18	-29	Eager to go to bed	*vs.* Resists going to bed
24	-21	Afraid of strangers	*vs.* Unafraid of strangers
33	-20	Seldom has temper tantrums	*vs.* Often has temper tantrums

Identification as D is tentative. Interpretation is impeded by the failure of any clear D pattern to emerge in the study with which the present results were statistically compared. However, loadings on a factor alternatively labeled Q^4 (Ergic tension) or D in the study of seven-year-olds correlate .37 with the present ones, and this suggests that a significant if not strikingly stable dimension which *may* represent the D prototype appears with reasonable consistency in young children.

FACTOR 5. SURGENCY-DESURGENCY (F)

39	52	Compliant	*vs.* Not compliant
3	-46	Generous	*vs.* Selfish
24	46	Unafraid of strangers	*vs.* Afraid of strangers
5	-34	Talkative	*vs.* Taciturn
6	30	Active, always on the go	*vs.* Inactive, Bored
26	-28	Bites fingernails	*vs.* Does not bite fingernails
1	27	Prefers younger playmates	*vs.* Prefers older playmates
42	-27	Affectionate	*vs.* Not affectionate
23	-26	Unafraid of animals	*vs.* Afraid of animals
37	-25	Cheerful	*vs.* Solemn

As has happened in a number of previous studies [1, 4] two surgency-like factors (see Factor 10 below) appeared in the present series. The sense of the factor now under discussion undeniably is that of F. Nearly all of the crucial components, happy-go-lucky, easy-going cheerfulness, generosity, and energetic activity, are represented, and the interpretative difficulty lies not so much in identifying this factor, which seems clearly to be F, as in the interpretation of Factor 10 to follow, and the two analogous factors in the study of seven-year-olds. Loading correlation between the present factor and the one finally designated as F for seven-year-olds is zero. The correlation with Factor 9 in the latter study, however, is .41, and this suggests reconsideration of the results of the previous study. Factor 9 in that investigation was ultimately labeled M (Autia), but the possibility of identification as F was seriously entertained, and the present findings suggest that such an alternative designation (F) might well have been chosen. Even if this were done, however, the

relationship between F and M would be somewhat obscure, and it will probably remain so until research is directed specifically toward its clarification.

FACTOR 6. PREMSIA VS. HARRIA (I)

36	47	Toilet trained early	*vs.* Toilet trained late
24	41	Unafraid of strangers	*vs.* Afraid of strangers
1	38	Prefers younger playmates	*vs.* Prefers older playmates
44	33	Not enuretic	*vs.* Enuretic
15	28	Has little physical endurance	*vs.* Has great physical endurance
7	25	Inquisitive about sex	*vs.* Not inquisitive about sex
2	24	Not destructive	*vs.* Destructive
34	−24	Talked early	*vs.* Talked late
23	−23	Unafraid of animals	*vs.* Afraid of animals
42	−21	Affectionate	*vs.* Shows little affection
40	−20	Daydreams excessively	*vs.* Seldom daydreams

The characteristic tenderness and sensitivity of I are manifested most clearly in affectionate qualities, tendencies toward daydreaming, and lack of physical toughness. Associations with early toilet training and rapid speech development are probably best explained by considering the latter to be part of the protective, attentive, and perhaps accelerational family influence which seems to form the background for the development of premsic tendencies. Correlation of loadings with those for I in seven-year-olds is only .23, but I appeared with relatively low variance in the latter study, and its identification there was somewhat uncertain.

FACTOR 7. OLDEST CHILD

31	56	Seldom has colds	*vs.* Has frequent colds
28	−41	Masturbates	*vs.* Does not masturbate
43	−41	Eats anything served	*vs.* Refuses to eat certain foods
4	36	Fearful	*vs.* Fearless
12	30	Often uses forbidden words	*vs.* Seldom uses forbidden words
5	27	Taciturn	*vs.* Talkative
32	26	Seldom has headaches	*vs.* Has frequent headaches
42	−24	Affectionate	*vs.* Shows little affection
6	−21	Inactive, Bored	*vs.* Active, Always on the go

There is very little in the list of salients above to suggest that the pattern is in any way associated with ordinal position among siblings. It is only the rather unimpressive correlation of .31 with the "oldest child" factor in Cattell and Coan's investigation [4] which supports our designative choice, and whether the factor is indeed related to the pressures and attentions associated with first sibling position can only be proposed as an hypothesis for future examination.

FACTOR 8. COASTHENIA (J)

18	−65	Eager to go to bed	*vs.* Resists going to bed
35	−60	Not domineering	*vs.* Domineering
11	29	Quiet in speech	*vs.* Loud in speech
5	27	Taciturn	*vs.* Talkative
34	25	Talked late	*vs.* Talked early
6	−22	Inactive, Bored	*vs.* Active, Always on the go

Interpretation as J is based principally on the passive, quiet, "neurasthenic" quality of the variables above. Notably absent from the pattern are the individualistic, obstructive tendencies so conspicuous in the adult representation of J. It seems possible that the revolt against early imposition of cultural pressure which such tendencies have been held to represent [1] is a rather late developmental phenomenon, and that the typical first response to such pressure is the relatively simpler one of passive acceptance. J was not identified in the study of seven-year-olds cited so frequently herein [4], but *two* I factors emerged, and it is interesting to note that loadings on the present factor correlate .51 with loadings on the second of these (Factor 14 in the study of seven-year-olds). Reinterpretation of that factor as J should be considered. In it, passive, anergic qualities were coupled with a tendency to complain, and it is conceivable that the latter represents an early form of revolt which has not appeared at all by the age of four, but which develops increasing prominence to become the obstructive individualism of the adult J pattern.

FACTOR 9. DOMINANCE VS. SUBMISSIVENESS (E)

38	70	Shows off	*vs.* Does not show off
41	51	Unhappy unless with others	*vs.* Prefers to be alone
36	30	Toilet trained early	*vs.* Toilet trained late
27	−28	Never sucks thumb	*vs.* Sucks thumb
26	−24	Bites fingernails	*vs.* Never bites fingernails
5	−23	Talkative	*vs.* Taciturn
4	21	Active, Always on the go	*vs.* Inactive, Bored
10	−20	Likes school	*vs.* Dislikes school
13	−20	Often sulks or pouts	*vs.* Seldom sulks or pouts

Very little over-all resemblance obtains between this representation of E and the similarly labeled factor in seven-year-olds. The loading correlation is only .11, and our interpretation is based solely on the vigorous, aggressive connotation of many of the salient variables. Apparent resemblance is closer to the Dominance factor as it appeared in peer ratings of 11-year-olds[5]. In that factor, as in this, talkativeness was fairly prominent, and support is again offered for the contention[1] that talkativeness in children is more a form of self-assertion than of pure sociability.

FACTOR 10. AUTIA VS. PRAXERNIA (M)

15	−53	Has great physical endurance	*vs.* Has little physical endurance
37	−44	Cheerful	*vs.* Solemn
6	36	Active, Always on the go	*vs.* Inactive, Bored
1	−35	Prefers older playmates	*vs.* Prefers younger playmates
42	−34	Affectionate	*vs.* Shows little affection
3	−29	Generous	*vs.* Selfish
9	25	Seldom complains	*vs.* Often complains
13	−23	Often sulks or pouts	*vs.* Seldom sulks or pouts
2	−22	Destructive	vs. Not destructive

In the discussion of Factor 5, it was noted that two surgency-like factors had appeared in this series. The combination of cheerfulness, activity, affectionateness, and generosity in the set of variables above suggest possible interpretation of this as a second surgency dimension, and consideration of this possibility should be retained through any future evaluations of the results. Our decision to call the factor M is supported by a loading correlation of .45 with loadings on the factor called Autia in the study of seven-year-olds and by a general resemblance in meaning to that factor. Imaginative tendencies (as opposed to practical-minded ones) and excesses of day-dreaming, however, appeared as the two most highly saturated variables in the seven-year-old representation of M, and the failure of either to appear here discourages any final assurance about our interpretation. The first of those variables (imaginative *vs.* practical-minded) could not emerge on any factor, as it was not included in the basic rating set, but the second could have appeared, and its failure to do so is puzzling. Resemblance to the adult pattern is not close; the unconventionality, eccentricity, and sensitively imaginative qualities which define the factor at maturity are nowhere in evidence at this early age. If this is indeed M, its most striking feature at age four is a general restlessness. By age six or seven, dreamy, imaginative tendencies have entered the constellation, but it is evidently not until much later that the eccentricity and Bohemianism of the adult pattern appear.

FACTOR 11. PROTENSION VS. INNER RELAXATION (L)

33	54	Has frequent temper tantrums	*vs.* Seldom has temper tantrums
7	−45	Not inquisitive about sex	*vs.* Inquisitive about sex
13	−41	Often sulks or pouts	*vs.* Seldom sulks or pouts
43	−30	Eats anything served	*vs.* Refuses to eat certain foods
8	24	Seldom cries	*vs.* Often cries
23	24	Afraid of animals	*vs.* Unafraid of animals
32	23	Seldom has headaches	*vs.* Has frequent headaches
44	−22	Enuretic	*vs.* Not enuretic
16	−20	Stutters	*vs.* Speaks fluently
26	20	Never bites fingernails	*vs.* Bites fingernails

Suspiciousness and proneness to jealousy are usually regarded as the principal markers for L. Neither was examined in this study, and their absence leads to the same kind of difficulty in interpretation which has plagued us throughout this report.

Tentative identification as L is based principally on the quality of tension implicit in a conjunction of tendencies to sulk and pout, but not to cry, and to explode occasionally in outbursts of temper. The readiness to employ projection in the relief of tension is not manifest, but again it is impossible to say whether this is a function of subject age or the failure to include any variables clearly relating to projective defense in the set examined. If our interpretation is correct, and this is in fact the early appearance of the protensive pattern, it seems possible that homosexual tendencies, alleged by psychoanalytic theorists to lie at the basis of paranoia, may make their initial appearance as a simple lack of normal curiosity about sex. This hypothesis, however, rests on such obviously tenuous assumptions that it can only be stated with limited force.

FACTOR 12. PARMIA VS. THRECTIA (H)

10	–58	Likes school	*vs.* Dislikes school
22	–52	Steals	*vs.* Does not steal
2	–36	Destructive	*vs.* Not destructive
18	–31	Eager to go to bed	*vs.* Resists going to bed
26	–24	Bites fingernails	*vs.* Does not bite fingernails
11	–23	Loud in speech	*vs.* Quiet in speech
6	20	Active, Always on the go	*vs.* Inactive, Bored

Correlation with loadings on H in seven-year-olds is .39. A major difference between the pattern above and its form at age seven and at maturity is the absence here of any clear evidence of sociability. Once again, this is most simply explained by the deficiency of relevant markers among the elemental variables. The only variable which is pertinent at all, *i.e.*, #41. Unhappy unless with others vs. Prefers to be alone, did not develop a sufficiently high loading (+.13) to appear in the list of salients above, but its direction, at least, is consistent with interpretation as H. At any rate, recent evidence[1] suggests that sociability is peripheral to the basic meaning of H, and that a fundamental immunity to threat lies at the core of Parmia. If so, and if truly we are dealing with H, it seems likely that early parmic expressions may often take the form of socially-oriented aggressive action (variables 22 and 2), performed not so much in the satisfaction of any deviant motives, but because the usual internal emotional controls are relatively feeble. Liking for school (variable 10) becomes comprehensible when one considers the meaning of school to a four-year-old child. Unlike his older counterpart, in whom liking for school is a multiply determined attitude, the four-year-old is likely to be concerned principally with whether this, his first prolonged exposure to the non-family world, is dangerous or not, and his liking or antipathy toward school would seem most probably to depend on his reaction to the threat which leaving home and going to school implies.

FACTOR 13. EGO STRENGTH VS. PRONENESS TO NEUROTICISM (C)

11	45	Quiet in speech	*vs.* Loud in speech
38	–36	Seldom shows off	*vs.* Shows off
27	30	Sucks thumb	*vs.* Never sucks thumb
23	–28	Unafraid of animals	*vs.* Afraid of animals
15	–26	Has great physical endurance	*vs.* Has little physical endurance
40	26	Seldom daydreams	*vs.* Daydreams excessively
26	24	Never bites fingernails	*vs.* Bites fingernails
33	–24	Seldom has temper tantrums	*vs.* Has frequent temper tantrums
5	23	Taciturn	*vs.* Talkative
18	–22	Eager to go to bed	*vs.* Resists going to bed

Curiously, this low variance factor is easier to interpret than many of the larger ones. Loading correlation with C in seven-year-olds is .54, and the character of the salients, with quietness, calmness, stability, and freedom from fear predominant, is thoroughly consistent with the chosen interpretation. The most striking discrepancy from the expected pattern lies in the appearance of thumb-sucking at the C(+) pole. In the absence of any clear evidence that such behavior is associated with later emotional disturbance, no explanation of its apparent reversal here seems required; the early occurrence of thumb-sucking may even imply a positive prognosis in respect to later adjustment.

FACTOR 14. CYCLOTHYMIA VS. SCHIZOTHYMIA (A)

9	46	Seldom complains	*vs.* Often complains
8	33	Seldom cries	*vs.* Often cries
3	-30	Generous	*vs.* Selfish
39	29	Compliant	*vs.* Not compliant
41	-28	Prefers to be alone	*vs.* Unhappy unless with others
26	-26	Bites fingernails	*vs.* Does not bite fingernails
19	-25	Over-achieves in school	*vs.* Under-achieves in school
4	-25	Fearless	*vs.* Fearful
16	-23	Stutters	*vs.* Speaks fluently
17	-22	Articulates poorly	*vs.* Articulates well

Loading correlation with A as defined in the study of seven-year-olds is .62. Like that factor, the present one shows a disinclination to complain and a general willingness to "take things as they come" as the outstanding features of A in young children. Various differences from the adult pattern are apparent in the variables with relatively low loadings, and of these the most striking is the occurrence of a preference for *solitary* activities (variable 41) at the A(+) pole. The recent relegation of sociability to a secondary position in the meaning of A [1] is thus supported, and current theoretical focus on flexibility as the dynamic core of the factor is further justified.

SUMMARY

Eighty children in a metropolitan university nursery school were rated by their parents in respect to 43 variables. Reports of teachers were obtained for an additional variable and the entire set intercorrelated. Fourteen factors were extracted, and after an initial orthogonal solution had been reached through use of the quartimax routine for electronic computer, two independent series of rotations to oblique simple structure were carried out. One of these entailed knowledge of item content at one point in the process; the other was "blind" throughout. The factors reported represent the best single solution to emerge from both kinds of rotational procedure. Factors were compared in the usual quasi-judgmental way with those isolated in earlier investigations [1], and were matched statistically with the factors reported in a previous study of parent ratings on six- and seven-year-old children [4]. Results further confirmed the now rather well-established proposition [3, 4, 5, 6, 7, 8, 9, 10] that personality structure in young children is not radically different from personality structure in older children and adults. Seven factors (G, K, O, I, H, C, and A) could be recognized with reasonable clarity and bore a statistically demonstrated resemblance to similarly identified factors in the study of six- and seven-year-olds. Five others (D, E, M, L, and a putative "oldest child" factor evidently unique to young children) presented a variety of interpretative problems, and only tentative identifications were made. In two cases (F and J), it was suggested that interpretations given in the report on six- and seven-year-olds be reconsidered in the light of new information gained in the present study.

In spite of the general resemblance between the early patterns reported herein and those found in more mature individuals, certain disparities occurred. Some of these probably represent systematic, age-related modification in personality expression, and a number of hypotheses were offered in reference to such apparent developmental changes.

REFERENCES

1. CATTELL, R. B. *Personality and motivation structure and measurement.* Yonkers, N. Y.: World, 1957.
2. CATTELL, R. B. and BAGGALEY, A. R. The salient variable similarity index— s —for factor matching. Urbana, Ill.: Lab. of Personality Assessment. Adv. Publ. No. 4, 1954.
3. CATTELL, R. B. and COAN, R. W. Child personality structure as revealed in teachers' behavior ratings. *J. clin. Psychol.*, 1957, *13*, 315-327.
4. CATTELL, R. B. and COAN, R. W. Personality factors in middle childhood as revealed in parents' ratings. *Child Develpm.*, 1957, *28*, 439-458.
5. CATTELL, R. B. and GRUEN, W. The personality factor structure of 11-year-old children in terms of behavior rating data. *J. clin. Psychol.*, 1953, *9*, 256-266.

6. CATTELL, R. B. and GRUEN, W. Primary personality factors in the questionnaire medium for children eleven to fourteen years old. *Educ. psychol. Measmt.*, 1954, *14*, 50-76.
7. CATTELL, R. B. and GRUEN, W. The primary personality factors in eleven-year-old children by objective tests. *J. Pers.*, 1955, *23*, 460-478.
8. CATTELL, R. B. and PETERSON, D. R. Personality structure in four- to five-year-olds: Functional unities operating in combined rating and life record observation. *Child Develpm.*, in press.
9. CATTELL, R. B and PETERSON, D. R. Personality structure in four- to five-year-olds by factoring observed (time sampled) behavior. *J. consult. Psychol.*, in press.
10. PETERSON, D. R. and CATTELL, R. B. Personality factors in nursery school children as revealed by teacher ratings. *J. educ. Psychol.*, in press.

PERSONALITY FACTORS IN THE READINESS TO EXPRESS AGGRESSION[1]

PHILIP WORCHEL

University of Texas

PROBLEM

A number of studies indicate the importance of adequacy, self-sufficiency, and adjustment in the susceptibility to frustration and the readiness to express direct aggression. Wright[9], French[3], Zander[10], and Pepitone and Reichling[7], find that a person who feels secure in a situation can "let off steam" by voicing his feelings about the situation. Doris and Sarason[2] find that high-anxious persons are most likely to blame themselves for failure. Self theory and experiments validating its predictions demonstrate that congruency between self and ideal is related to feelings of security, anxiety, and susceptibility to threat. A person with a large discrepancy in his self-ideal is chronically frustrated in attaining his ego-ideal or personality aspirations. He has strong feelings of self-devaluation, helplessness, and insecurity. Such persons are more likely to blame themselves for not attaining their "ideal" rather than the instigators of frustration.

It is therefore predicted that *Ss* with low self-ideal discrepancies will be more likely to express direct aggression towards authority figures than will those *Ss* with high discrepancies. Also since Dollard *et al*[1] state that there should be a strong tendency for inhibited aggression to be displaced, then *Ss* who are prevented or inhibited in their expression of aggression, should have a greater tendency to displace their aggression than those with high self-ideal discrepancy.

EXPERIMENT I

Instruments. An "intelligence" test, used in a previous study for the arousal of hostility, was administered to all *Ss* under motivating instructions. In order to make the test as realistic as possible, the first two sub-tests and two others interspersed in the middle consisted of relatively easy items which almost all *Ss* were able to complete in the time limits.

To secure a rating of the degree of hostility experienced by *Ss* towards the instigator of frustration, an attitude scale entitled "Rating of Test Administration" was constructed. This scale consisted of 25 items describing in both positive and negative terms the reactions of *Ss* towards the examiner (instigator) in administering the intelligence test. The *Ss* were requested to place a check mark in one of the five blanks next to each item which best described his agreement from "Yes, I strongly agree" to "No, I strongly disagree." The total score was the sum of the ratings on each item from 1 to 5 with the higher rating always indicating negative (or hostile) feelings. The reliability of the scale was determined by correlating the odd *vs.* even items for 103 *Ss*. The corrected split-half product-moment correlation yielded an *r* of .92.

[1]One phase of this study was aided by the Behavioral Science Division, Air Force Office of Scientific Research, ARDC, under project AF 49(638)-33.

To secure a measure of displacement of aggression following frustration to persons similar to the instigator, a rating scale entitled "Attitude Towards Instructors" was constructed. This scale was similar to the "Rating of Test Administration". It contained 26 items describing instructional ability in both positive and negative terms. The score was the sum of the ratings from 1 to 5 on each item where the higher score indicated the degree of negative feelings expressed by *Ss* towards instructors in general. The split-half product moment correlation for 103 *Ss* was .82, indicating fairly high internal consistency.

The Self-Activity Inventory (SAI), developed by the author for the USAF[2], was used as a measure of the self-ideal discrepancy. The total discrepancy score (SI) was obtained by summing the absolute discrepancies on each item between self and ideal.

Subjects and Procedure. For the present experiment, four classes consisting of 104 *Ss*, 43 males and 61 females, in elementary educational psychology were selected. All four classes received the same treatment. The *E* mentioned that intelligence and personality tests are administered routinely to all classes in psychology and educational psychology in order to obtain data for personal counseling, for purposes of predicting academic success, and for recommendations for graduate work and to prospective employers.

The intelligence test was then administered with the instruction that each subtest was timed but that sufficient time had been allowed for most of the students to complete the tests. As the students worked, *E* walked around noting the students' progress, urging faster work, belittling their efforts, and unfavorably comparing their poor performance to that of other classes. He interrupted frequently by urging them to skip over the ones they could not do and to try the easier ones first.

At the completion of the intelligence test the class was then informed that it was important to obtain their reactions to the test administration as such evaluations help in the revision of tests. The attitude scale on "Rating of Test Administration" was then distributed. The students were informed to fill out the blanks at the top of the page. At this point a randomized procedure was used to split the class into experimental (identification) and control (no-identification) groups. On one-half of the papers information was requested as to name, marital status, class section, degree expected, major subject; and on the other half of the papers, instructions at the top of the page stated, "Do not write your name. We want this evaluation to be anonymous so be as frank as you can. Just fill in the few blanks below for general data." The information requested was degree expected, major subject, and marital status. This information was requested so as to give the appearance that all *Ss* were filling out the same blanks. Instructions were then given to place a check mark in the blank at the right which best described their opinion.

At the conclusion of the attitude scale, *Ss* were told to leave their papers on the desk while they were taking the next test. The "Attitude toward Instructors" was then given out with the instructions that no name or any other data was to be placed on the scale as the ratings were to be anonymous. At the completion of the Scale, they were instructed to leave the sheets on the desk.

For the Self-Activity Inventory the *Ss* were requested to read the instructions about encircling the number in the Self column (I) and then in the Ideal column (II). When they finished the Inventory they could leave the room leaving all their papers on their desk. The *E* then picked up the papers keeping them all together so that he could identify the attitude scales (those without names) by the names on the Self-Activity Inventory. After all groups had been tested, *E* returned to each class and explained the "bogus" intelligence test and the nature of the experiment.

In order to test the hypotheses concerning the effects of Self-Ideal discrepancy on the inhibition and displacement of aggression, *Ss* were randomly selected to fill a

[2]The development of the Self-Activity Inventory (SAI) was supported in part by funds provided under Contract AF 18(600)-913 with the USAF School of Aviation Medicine, Randolph Field.

2 x 2 x 2 factorial design, that is, two degrees of identification (no name *vs.* name on the "Rating of Test Administration"), sex, and self-ideal discrepancy (high *vs.* low). The smallest number of *Ss* falling in the above categories were eight *Ss* for the "name-male-high SI" group. Of the other seven cells which contained from 9-12 *Ss*, eight *Ss* were randomly selected from each category to complete the 2 x 2 x 2 factorial design. Thus a total of 64 *Ss* were selected for the analysis.

Results

Table 1 shows the results of the analysis of variance of the scores on the attitude towards the instigator of aggression (test-administrator). The F-ratios for the variables of identification (no-name *vs.* name) and SI are significant beyond the .01

Table 1. Analysis of Variance of the Scores on the Attitude Towards Instigator and Attitude Towards Instructors

Source of Variation	df	Instigator Mean Square	F	Instructor Mean Square	F
Identification	1	3093.8	14.88**	20.3	——
Sex	1	6.5	——	451.6	3.14
Self-Ideal	1	1113.5	5.36*	126.6	——
Ident. x Sex	1	347.2	1.67	9.0	——
Ident. x Self-Ideal	1	236.7	1.14	6.2	——
Sex x Self-Ideal	1	15.4	——	18.1	——
Sex x Self-Ideal x Ident.	1	129.1	——	182.2	1.27
Within groups	56	207.9		144.0	
Total	63				

*Significant at the .05 level
**Significant at the .01 level

and .05 levels respectively. An analysis of the results on these two variables indicates that *Ss* under the "no-name" condition gave a significantly higher mean negative evaluation (75.93 ± 14.53) of the instigator than *Ss* under the "name" condition (62.03 ± 14.51). In other words, identifying one's evaluations of the instigator inhibited *Ss* on the average in the expression of negative feelings. This confirms the contention of the Yale group that the strength of the inhibition of any act of aggression towards the instigator varies with the amount of punishment anticipated as a consequence of that act. This assumes, of course, that "identification" results in a greater expectation of punishment to that individual, or that the anonymous condition minimizes the expectation of punishment since the person could not be identified.

Regarding SI discrepancy, those with high SI scores, regardless of whether they identified themselves or not, expressed significantly less hostile feelings on the average (64.8 ± 16.32) towards the instigator than those with low SI scores (73.2 ± 14.75), confirming the first hypothesis.

It was predicted also that the *Ss* who were prevented or inhibited in the direct expression of aggression ("name" condition) would displace their aggression towards similar stimuli. Table 1 shows the results of the analysis of variance of the scores on the "Attitude Towards Instructors". None of the F-ratios attained the .05 level of significance. The means for both the high and low SI groups (65.4 and 62.6) and for the "name" and "no-name" (63.5 and 64.6) conditions are almost identical. Thus the prediction of displacement was not confirmed. It is interesting to note, however, that the F-ratio for sex was almost significant at the .05 level. The females had a greater tendency to displace their aggression than males.

One reason for the negative results may be that the hypothesis of displacement as formulated by the Yale group is not valid. It is also possible, however, that the "name" group had already expressed some aggression towards the test administrator thus minimizing the tendency to displace. At the same time, there may have been *Ss* in the "no-name" group who may not have fully expressed direct aggression and therefore may have displaced their feelings to some extent to the "instructors". To

test this hypothesis conditions were varied in the following experiment so that different degrees of expression of direct aggression were permitted.

Experiment II

To control the degree of communication of hostile feelings, four groups of *Ss* were selected. The control group (I) consisted of 58 *Ss* randomly selected from four other educational psychology classes. The rating scale on "Attitude towards Instructors" was administered without previous frustration. Instructions were the same as in the first experiment, and ratings were anonymous. Group II consisted of all 47 *Ss* participating in the first experiment in the "no-name" condition thus providing for maximum freedom to express direct aggression towards the test administrator. Group III consisted of all 57 *Ss* under the "name" condition of the first experiment who were significantly inhibited from expressing direct aggression. The fourth group was not allowed to communicate any hostile feelings about the instigator. It consisted of 46 *Ss* from the remaining two educational psychology classes (Group IV) who were subjected to the frustrating experience (the intelligence test), as in the previous experiment, but were immediately requested to rate instructors. The data for all four groups on the ratings of instructors are given in Table 2. As can be

Table 2. The Means and SDs of the Ratings on Attitudes Towards Instructors for All Four Groups

	GROUP			
	I	II	III	IV
Mean	66.43	65.13	64.54	67.22
SD	7.88	11.52	10.58	8.63
N	58	47	57	46

seen from this table, the means are almost identical. The F-ratio is not significant at the .05 level, and thus we can conclude that there is no signficant difference among the four groups in the expression of negative attitude towards instructors in general. We can conclude therefore that the displacement hypothesis proposed by the Yale group is not valid under the conditions of the present experiment.

Discussion

The present study predicted that individuals with a high discrepancy between self- and ideal-concepts would have a lower tendency to instigate direct aggression than *Ss* with low discrepancy. The results confirmed the hypothesis. Under conditions encouraging displacement of hostility, however, *Ss* who were inhibited in the expression of direct aggression ("name" group) did not "attack" instructors. It is possible, however, the "instructors" represent a class of persons about whom *Ss* had crystallized or overlearned a certain set of responses prior to the experiment and therefore would be less likely to displace aggression towards them. It is also possible that the rating scale contained statements of fact which may be difficult to deny and therefore less sensitive as a measure of displacement particularly in intelligent *Ss*. For example, "instructors start their lectures promptly" may be difficult to deny whereas such ambiguous statements as "lectures do not appear well-planned" or "I always have opportunity for fair treatment by instructors" may be more easily affected. Not only may such variables as self-ideal discrepancy affect the communication of aggression, but the intensity of hostility and the level of intelligence may also influence the choice of the stimulus in displacing aggression.

There have been a sufficient number of experiments[4, 5, 9] which have failed to find displacement when predicted and which therefore cast doubt upon the validity of the displacement hypothesis. Dollard, *et al*[1] state that aggression may be directed at the object which is perceived as causing frustration or it may be displaced to some altogether *innocent*[3] source. Perhaps, the results of the present experiments and the

[3]Italics mine.

ones which failed to find displacement might suggest that the object of displacement must also be frustrating. Inhibitions of direct aggression may lower the threshold of frustration-tolerance so that any object might be the target for displaced aggression if it is also frustrating. The additional frustration, however, might be too weak by itself to arouse an aggressive response but, falling upon a lowered threshold of frustration-tolerance induced by the first frustration, now becomes "super-threshold." Zawadzki [11] in criticizing the inadequacies of the "scapegoat" and "well-earned reputation" theories holds a similar position in his convergence theory of prejudice in maintaining that both internal factors (drive) and the objective characteristics of the stimulus and object of the reactions must be taken into account. The results of the present experiment might have been different if "instructors" had acted, even mildly, as frustrating agents.

Summary

Hypotheses relating self-ideal discrepancy and the readiness to express direct aggression following frustration predicted that persons with low self-ideal discrepancy would be more likely to express direct aggression towards authority figures than those with high discrepancy. It was further predicted from the displacement hypothesis that *Ss* inhibited in direct expression of aggression would have a greater tendency to displace their aggression towards a figure similar to the instigator. A 2 x 2 x 2 factorial design involving sex, self-ideal discrepancy, and identification consisting of a total of 64 *Ss* was used. An attitude scale evaluating hostility aroused by frustrating performance on an "intelligence" test was administered with one half of the group identifying themselves by name, class, sex, etc. and the other half, without any personal identification. Following this, a test for displacement of aggression consisting of a critical evaluation of instructors in general was filled out anonymously by all *Ss*. A second experiment was designed to test explanations suggested for the negative results on displacement. One group was administered only the attitude scale towards instructors and another group was subjected to frustration followed immediately by the displacement situation. The results show that:

1. Under the "no-name" condition, *Ss* expressed significantly greater negative feelings towards the instigator than under the "name" condition.

2. *Ss* with low self-ideal discrepancy expressed significantly greater direct aggression than those with high self-ideal discrepancy.

3. There was no significant difference between the sexes in their expression of direct aggression following frustration.

4. There was no evidence for displacement of aggression towards instructors under any of the conditions in the experiment.

References

1. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., and Sears, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
2. Doris, J. and Sarason, S. Test anxiety and blame assignment in a failure situation. *J. abnorm. soc. Psychol.*, 1955, *50*, 335-338.
3. French, J. R. P., Jr. The disruption and cohesion of groups. *J. abnorm. soc. Psychol.*, 1941, *36*, 361-377.
4. Graham, F. K., Charwat, W. A., Honig, A. S. and Weltz, P. C. Aggression as a function of the attack and the attacker. *J. abnorm. soc. Psychol.*, 1951, *46*, 512-520.
5. Lindzey, G. An experimental examination of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.*, 1950, *45*, 296-309.
6. Miller, N. E. and Bugelski, R. Minor studies of aggression. II. The influence of frustrations imposed by the in-group on attitudes expressed toward out-groups. *J. Psychol.*, 1948, *25*, 437-442.
7. Pepitone, A. and Reichling, G. Group cohesiveness and the expression of hostility. *Hum. Relat.*, 1955, *8*, 327-337.
8. Stagner, R. and Congdon, C. S. Another failure to demonstrate displacement of aggression. *J. abnorm. soc. Psychol.*, 1955, *51*, 695-696.
9. Wright, M. E. Constructiveness of play as affected by group organization and frustration. *Charact. Pers.*, 1942, *11*, 40-49.
10. Zander, A. F. A study of experimental frustration. *Psychol. Monogr.*, 1944, *56*, No. 3, 1-38.
11. Zawadzki, B. Limitations of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.*, 1948, *43*, 127-141.

THE RELATIONSHIP BETWEEN THE SELF CONCEPT AND THE IDEAL-SELF CONCEPT AS A MEASURE OF ADJUSTMENT[1]

IRWIN D. NAHINSKY

Oregon State Hospital, Salem, Oregon

Rogers *et al.*[2] emphasized that various outcomes of psychotherapy are reflected in changes in the client's perceptions. Butler and Haigh[1] postulated that the degree of discrepancy between the self concept and the ideal-self concept is a measure of self esteem, showing that this discrepancy decreases in "client-centered" or non-directive therapy. Presumably, this indicates an increase in self acceptance, one of the ostensible goals in therapy. It would seem desirable to determine if the discrepancy or congruence between self concept and ideal-self concept manifests itself in acceptance or rejection of certain roles in the external world. The present study attempts to investigate the relationship between self-ideal discrepancies and adjustment in a specific environment. The Q-technique developed by Stephenson[3] was utilized in this study.

The Experiment

Sample. The sample consisted of 74 junior officers leaving the Navy, referred to as the "non-career group", and 35 junior officers who had recently been accepted for retention in the Regular Navy, referred to as the "career group". The officers leaving the Navy were contacted at the separation centers in San Diego and Long Beach. The retained officers were contacted in the central and southern California area. It was not possible to sample in a strictly random fashion. However, it was possible to get the data from nearly all of the non-career officers being processed out on days during which the data were collected, and from nearly all officers recently transferring to the Regular Navy who were available in the area. Therefore, it is unlikely that the sample was biased by self-selection. Each group also had a fairly proportional representation of the main officer specialty areas and sources of commission.

Procedure. A forced-choice inventory of 100 statements was constructed to assess the perceptions of officers in the sample. Each officer in both groups was asked to sort 100 cards into a forced normal distribution three times. One sorting was to describe himself, another a "typical career Navy officer", and another an "ideal career Navy officer". On each card was typed the statment of an attribute such as: "He (or I) can do anything his (or my) men can do" or "He is (or I am) patient". The items for the inventory were developed largely on the basis of preliminary interviews of junior officers leaving the Navy.

The cards were sorted into seven categories for each of the three descriptions. Category seven was reserved for items "most like" the person or stereotype being described, and category one was reserved for items "least like" the person or stereotype being described. The other five categories were used for statements that ranged between the two extremes. The sorting was done on a board which contained seven columns of slots, one column for each category. The number of slots corresponded to the number of cards necessary for each category. Categories one and seven had one slot each; catagories two and six had six slots each; categories three and five had 24 slots each; and category four had 38 slots. This obviated the necessity of a subject's counting the number of cards he had sorted, and permitted simultaneous inspection of all cards sorted at anytime during the sorting. The cards were shuffled after each

---

[1]The research was done at the U. S. Naval Personnel Research Field Activity of San Diego. The opinions expressed here do not necessarily reflect the opinions of this organization or of the Bureau of Naval Personnel. The writer acknowledges the contribution of Richard E. Farson of San Diego, who assisted in developing and selecting the items for the measuring instrument used in this study.

sorting in order to eliminate bias that might arise from the previous order of the cards.

To control for the possible effects which might result from the sequence in which the three concepts were sorted, a complete counter-balancing was effected using the six possible sequences of sorting for the three concepts. It was not feasible to have an equal number of subjects for each of the six sequences, because it was necessary to collect the data from different locations. However, a chi-square test reveals that the distribution of subjects among the six conditions does not differ significantly from an equal frequency of subjects for all of the conditions. This is true for both groups in the sample.

Results

Comparisons of the Two Groups for Correlations between Concepts. For each officer three correlations were computed, one between the self and typical officer values that he had assigned, one between the self and ideal officer values, and one between the ideal officer and typical officer values.

Analyses of variance were made between the two groups of officers for each of the three correlations. The variable used for analysis was the z-transformation of the correlations. The correlations corresponding to the mean z-values for each pair of concepts for each of the two groups are shown in Table 1. The percentage of non-

Table 1. Mean Product-Moment Correlations for Three Pairs of Concepts for the Career And Non-Career Groups

Correlation	Non-career officers (N = 74)	Career officers (N = 35)	F ratio	Non-career officers exceeding the mean of the career group	
				N	Percentage
Ideal officer concept—self concept	.37	.54	21.20**	14	19%
Typical officer concept—ideal officer concept	.29	.36	1.86	24	32%
Typical officer concept—self concept	.21	.34	5.86*	25	34%

Note: Means obtained by use of the z-transformation
*Significant at the .05 level
**Significant at the .01 level

career officers that exceed the mean correlation of the career group is also shown for each of the correlations. L-tests used to test the homogeneity of variance between the two groups yielded non-significant results for each of the three correlations. The ideal officer—self correlation shows a very significant discrimination between the two groups with an F-ratio of 21.20, which is significant at the .01 level. A mean correlation of .54 was found for the career group as compared with .37 for the non-career group. The typical officer-self correlation yielded an F-ratio of 5.86, which is significant at the .05 level, again in favor of the career officers. The mean correlation for the career group is .34, and for the non-career group it is .21. The ideal officer-typical officer correlation yielded an F-ratio of 1.86, which is not significant. The average correlation for the career group is .36, and for the non-career group .29.

Thus in all cases the correlations between the various concepts were higher on the average for officers remaining in the Navy than for officers leaving the Navy. The fact that the self-ideal discrepancy showed the most significant discrimination indicates that the feeling of inadequacy or not "measuring up" may be the most important correlate of adjustment. However, the significant discrimination for the self-typical discrepancy indicates that a feeling of "fitting in" is an important correlate of situational adjustment.

Analysis of Variance for the Interaction of Items with Group and with Concept. An analysis of variance of item values, which is shown in Table 2, was done to assist in

TABLE 2. ANALYSIS OF VARIANCE FOR ITEM SCORES ON A FORCED-CHOICE INVENTORY USING THE DIMENSIONS OF SORTS, SUBJECTS AND OFFICER GROUPS (N = 109)

Source of variance	Degrees of freedom	Variance	F ratio
Between items	100	62.21*	6.93**
Within items			
Between concepts:			
Item-by-concept interaction	200	8.98	12.72**
Item-by-concept-by-group interaction	200	1.18	1.67**
Error	21400	.71	
Between concepts total	21800		
Between groups:			
Group-by-item interaction	100	2.10	1.64**
Error	10700	1.28	
Between group total	10800		
Within items total	32600		
Grand total	32700		

*The item-by-concept interaction was used as the error term for the between items test.

**Significant at the .01 level

clarifying the relationship between the correlations. An analysis of the group-by-concept-by-item interaction resulted in an F-value of 1.67, significant at the .01 level, which shows a significant tendency for the two groups each to assign items different values for the different concepts. The group-by-item interaction yielded an F-value of 1.64, significant at the .01 level, which indicates that the two groups gave significantly different values to the items, on the average, over all concepts. The between-items variance, which was by far the largest, yielded an F-value of 6.93, significant at the .01 level, which indicates a generalized tendency for certain items to be assigned larger values than others regardless of the concept in which they are sorted. However, the concept-by-items interaction yielded an F-value of 12.72 which is significant at the .01 level. Thus there is a tendency for the particular concept to influence the value given to an item despite any general tendency that exists to assign certain items a higher value no matter where they appear.

An inspection of the individual items for which this between concept difference appears reveals that for 63 of the 100 items the between-concept F-ratio is significant at the .01 level, and for 14 of the 100 items the F-value is significant at the .05 level.

An analysis of the group-by-concept interaction for the individual items reveals only 4 items with F-values significant at the .05 level, and an additional two items with F-values significant at the .01 level, which is about what chance would predict. Thus the pattern of assigning values to particular items between the different concepts does not differ significantly between the two groups. This shows that the differences found between the two groups for the correlations are independent of the manner in which particular items were sorted.

Analysis of Individual Item Discriminations Between Groups. Despite the fact that correlations were not influenced by differential item values for the two groups, a total of 26 discriminations between the two officer groups were found for various items over all three concepts. Thirteen of these discriminations were within the self concept. Table 3 shows the items for all of the three concepts that discriminated be-

TABLE 3. ITEMS THAT DISCRIMINATE SIGNIFICANTLY CAREER FROM NON-CAREER OFFICERS

	Officers leaving the Navy		Officers remaining in the Navy		
	Means	S.D.	Means	S.D.	*F*
Self Concept Items					
6. I take special pride in my uniform	4.20	.94	4.83	.86	11.20**
20. I enjoy being by myself	3.80	.90	3.11	1.28	8.32**
21. It is important to me to be neat and well groomed	4.30	.82	4.66	.73	4.89*
23. I have a good speaking voice	4.18	.80	3.83	.86	4.27*
24. I feel that I deserve the credit for the accomplishment of my men	2.91	1.14	2.37	1.00	5.64*
28. Family life is very important to me	4.93	1.33	4.34	1.45	4.41*
32. I am very friendly with my men	4.06	.93	3.40	.74	13.98**
38. My men are more important than any mission	3.80	1.01	3.17	.89	9.87**
48. I identify very much with the Navy	3.65	1.31	4.60	1.12	13.76**
50. Service to my country is foremost in my values	3.86	.93	4.77	1.21	18.52**
54. I am self critical	4.45	.89	4.03	1.04	4.65*
68. I love my work	3.72	1.15	4.80	1.02	22.51**
74. I am flexible	4.35	.88	4.80	.83	6.36*
Typical Officer Concept Items					
1. He enjoys cultural activities	3.57	1.11	4.06	1.19	4.41*
28. Family life is very important to him	3.99	1.33	4.83	1.12	10.48**
63. He is very attentive to small details	3.82	.96	3.34	.97	5.98*
79. He does not question his superiors	4.31	1.28	3.71	1.10	5.62*
89. He is conscious of his status	4.93	1.20	4.43	.88	4.92*
98. He is unswerving in his attitudes	3.62	1.10	3.17	.86	4.54*
Ideal Officer Concept Items					
20. He enjoys being by himself	3.15	1.09	2.66	1.28	7.38**
24. He feels that he deserves the credit for the accomplishment of his men	2.78	1.04	2.29	.99	5.65*
28. Family life is very important to him	3.65	1.03	4.14	.88	6.02*
36. He is understanding	4.50	.71	4.80	.76	4.08*
40. Doing a good job is satisfaction enough for him	4.08	.99	3.51	1.12	7.15**
72. He is strict	4.08	.92	3.63	.77	6.37*
78. He is straightforward in all his dealings	4.66	.85	5.03	.92	4.19*

*Significant at the .05 level
**Significant at the .01 level

tween the two groups at least at the .05 level. Ten of these items discriminated at the .01 level. Within no concept is the number of significant items significantly greater than would be expected by random sampling error. However, despite the relatively small number of discriminating items, an examination of the F-ratios for the interaction between items and career intention groups for each concept reveals interaction significant at the .01 level for both the typical officer concept and the self concept, and an interaction significant at the .05 level for the ideal officer concept. Thus there was a significant tendency for each career intention group to picture oneself, the typical career Navy officer, and the ideal career Navy officer differently, although this tendency showed up strongly for relatively few individual items.

A number of the items that do discriminate between the career and non-career groups do not bear a relationship, which is very obvious, to qualities that one might expect to distinguish superior officers from inferior officers. For example, for the self-concept, discriminating items such as: "I enjoy being by myself," "I am self critical," and "It is important to me to be neat and well groomed," are not any more obviously related to officer proficiency than to other walks of life. For the ideal officer concept, the same can be said for items such as: "He is straightforward in all of his dealings," and "He enjoys being by himself." For the typical officer concept the items, "He enjoys cultural activities," and "He is unswerving in his attitudes," are examples of this tendency. There also appears to be some tendency for non-career officers to have more rigorous expectations than have career officers of the qualities of a superior Navy officer. For the ideal officer concept, items such as: "Doing a good job is satisfaction enough for him," and "He is strict," which have values higher for the non-career group than for the career group, illustrate this.

We have seen that the discrimination between the two groups in terms of self-ideal and self-typical discrepancies was strong despite only chance discrimination in terms of individual items. Also the differences between the groups in the correlations are independent of the way individual items were sorted. The author believes that this enhances the argument for the validity of the self-ideal discrepancy or the self-typical discrepancy as a measure of adjustment. The discriminations that these two measures made between individuals at different levels of adjustment or satisfaction make the measures seem even more useful when they are independent of specific stereotypes and idiosyncratic perception.

Summary

Two groups of Navy officers, one composed of officers choosing a Navy career, and the other of officers deciding against a Navy career sorted 100 statements, using the Q-sort. One sorting was to describe themselves, another to describe a typical career Navy officer, and another to describe an ideal career Navy officer. For each officer, correlations were computed between item values for each of the three pairings of descriptions. An analysis of variance of item values was also done to compare the groups. The data indicate the following conclusions:

1. Officers leaving the Navy picture themselves as being less like the typical career Navy officer and the ideal career Navy officer than officers staying in the Navy picture themselves.

2. These lower inter-concept relationships for the non-career group are not related to any differences between the two groups in the way individuals picture themselves, the typical career Navy officer, or the ideal career Navy officer.

3. The differences between the way the two groups pictured themselves, the typical career Navy officer, and the ideal career Navy officer were all significant.

4. The Q-technique proved useful in discriminating between two groups at different levels of adjustment to a particular situation.

References

1. Butler, J. M. and Haigh, G. V. Changes in the relation between self-concepts and ideal concepts consequent upon client centered counseling. In C. R. Rogers, and Rosalind F. Dymond (Eds.), *Psychotherapy and personality change.* Chicago: University of Chicago Press, 1954, pp. 55-75.
2. Rogers, C. R. and Dymond, Rosalind F. (Eds.). *Psychotherapy and personality change.* Chicago: University of Chicago Press, 1954.
3. Stephenson, W. *The study of behavior: Q-technique and its methodology.* Chicago: University of Chicago Press, 1953.

THE SELF CONCEPT IN PARANOID SCHIZOPHRENIA

ARTHUR H. ROGERS

Jacksonville (Ill.) State Hospital

PROBLEM

Friedman[2] found no significant difference between normals and paranoid schizophrenics in the comparison of their congruence of self concept and ideal self concept. Chase[1], on the other hand, found that psychotics in general and paranoid schizophrenics in particular had a significantly lower correlation between their self concept and ideal self concept than a group of normals. The present study approaches the self concept in paranoid schizophrenia with a new method for exploring the phenomenal world, a global apperception technique.

METHOD

Both Friedman[2] and Chase[1] used Q-sort data in their studies. The final correlation between the self concept and the ideal self concept in a Q-sort is at a relatively distant remove from the *S* than it is if the *S* can more directly compute the correlation himself in some way. A method in which an overall appraisal of the overlap of the self concept and ideal self concept is made directly by the *S* should be open to more defensive distortion, but such a variable is of interest precisely because it is the hypothesis of this study that paranoid schizophrenics show more defensive distortion than normals.

The apparatus used consists of two four-inch squares of colored glass set in a rectangular wooden frame, six inches high, sixteen inches wide, and one inch thick, mounted on legs. Facing the *S* on his left is a blue square of glass permanently affixed in a groove two inches from the left inner edge of the frame. Facing the *S* on his right is a red square of glass two inches from the right inner edge of the frame. The red square slides to the left in a groove directly behind the blue square. Overlapping the two squares produces a violet color that can be distinguished from the red and blue colors. Behind the blue square, out of the *S*'s sight, is a metric scale (10 cm long) which *E* can use to record the degree of overlap produced by the *S*.

Ss were asked to imagine that the red square represented the self ("the kind of person you are") and that the blue square represented the ideal self ("the kind of person you would like to be"). They were instructed to slide the red square to indicate the overlap, if any, between the two concepts. All *Ss* were requested to repeat what each of the squares represented, to make sure they understood the task required of them.

Ss were 30 patients diagnosed as schizophrenic reaction, paranoid type, and 30 psychiatric aides, all from Jacksonville State Hospital, Jacksonville, Illinois. The patients were examined within three months of their admission to the hospital. The patient group consisted of 17 women and 13 men, while the aide group consisted of 21 women and 9 men. Ages ranged from 20 to 64 in the patient group, with a mean of 38.67, and from 24 to 58 in the aide group, with a mean of 40.97.

RESULTS

Table 1 summarizes the results. As in Chase's study[1], a statistically significant difference was found between the psychotic sample and the normal sample.

TABLE 1. MEAN OVERLAP IN CENTIMETERS

Group	Mean	S.D.	Diff.	*t*	*p*
Patients	8.95	1.95	2.06	3.89	$<.001$
Aides	6.89	2.13			

The direction of this difference is opposed to Chase's finding, however. Chase found normals to have a significantly higher self-ideal self correlation than psychotics. In the present study the paranoid schizophrenics were found to have a significantly higher self-ideal self congruence than the normal group.

Discussion

In discussing the factors that may have led to the discrepancy between his findings and those of Friedman[2], Chase mentioned item content, and chronicity and severity of disturbance as follows: "It may be significant that Friedman's items can be said to be more pathologically oriented than those of the present study. Friedman has suggested that sampling differences may have led to the contradictory findings[1, p. 497]."

Two additional points deserve mention. First of all, since the method used in this study differs from a Q-sort, a more global level of personality involving greater defensiveness may have been tapped. Future comparisons of global apperceptions and the Q-sort method with the same *Ss* should help clarify the level of personality being explored by each technique. Secondly, the use of the diagnostic classification of paranoid schizophrenia is a confounded one, since the nonexistent classification of schizophrenic paranoid (primarily a paranoid psychosis with secondary schizophrenic features) would have described some of the *Ss* in the patient group more accurately than paranoid schizophrenia (primarily a schizophrenic psychosis with secondary paranoid features). Utilization of a more precise, reliable classification scheme, such as the one proposed by Wittman and Sheldon[3] could obviate the difficulties inherent in the classification system presently in use.

Summary

Using a global apperception technique, a new method for exploring the phenomenal world, an investigation was conducted of the congruence between the self concept and the ideal self concept in 30 paranoid schizophrenics and 30 normals. In contrast to two previous studies which found, using Q-sorts, no significant difference in one case and a significantly higher self-ideal self correlation in normals in the other, the present results showed that the paranoid schizophrenic group attained a significantly higher congruence between the self concept and the ideal self concept than the normal group. The results were discussed in terms of the method employed, with the global apperception technique presumed to be tapping a gross level of personality, where defensiveness could easily be employed. The need for a diagnostic category of schizophrenic paranoid to supplement that of paranoid schizophrenic was suggested.

References

1. Chase, P. H. Self concepts in adjusted and maladjusted hospital patients. *J. consult. Psychol.*, 1957, *21*, 495-497.
2. Friedman, I. Phenomenal, ideal, and projected conceptions of self. *J. abnorm. soc. Psychol.*, 1955, *51*, 611-615.
3. Wittman, Phyllis, and Sheldon, W. A proposed classification of psychotic behavior reactions. *Amer. J. Psychiat.*, 1948, *105*, 124-128.

THE PAIN APPERCEPTION TEST: PSYCHOLOGICAL CORRELATES OF PAIN PERCEPTION[1]

DONALD V. PETROVICH[2]

Veterans Administration Hospital
Jefferson Barracks, Missouri

PROBLEM

Reaction to pain has been recognized as something more than the automatic transmission of neural impulses. There is increasing awareness that emotional, experiential and situational factors are important in understanding individual responses to painful stimuli. The Pain Apperception Test [9] resulted from an extensive research project dealing with the psychological correlates of pain perception. Initial experimentation with this test indicates its suitability for investigating the psychological dimensions of pain response; and the present paper is intended to present some of the findings of the first experiment with this instrument.

Evidence from psychophysical studies indicates that neurotic and anxious subjects have lower pain thresholds than normals [1, 11]; and many writers have made the clinical observation that tense and neurotic people overreact to pain. Experimental results and clinical observation [6, 7] also indicate that anticipation is an important variable in pain reactions. The literature discloses no real indication of differences to be expected in comparing the reaction to pain which is self-inflicted with that which is inflicted by another person, although the study by Geers [5] suggests that control of the painful (electrical) stimulus by the subject may be an important factor in his reaction thereto. Geers also found no significant differences in cutaneous pain thresholds as related to subjective opinion regarding amount of pain experienced in one's lifetime or to one's ability to withstand pain as compared to other people.

At this initial stage of projective research into pain, our concern is with pain apperception correlates and the major purpose of this study was to investigate the consistency and uniqueness of the individual's pain apperception with other psychological attributes. The question is whether the perception of pain in the self as projected to stimulus pictures is a unique experience in the psychology of the individual or, on the other hand, this perception shows a relationship to his self-report of physical and psychological complaints associated with neuroticism, his self-report of manifest anxiety, and his perception of pain in others as projected in others (comparative self-evaluation). The variables of "anticipation" and pain instigation (self- and other-inflicted) were also investigated.

METHOD

Subjects. One hundred white, male patients of the Veterans Administration Hospital, St. Louis, Missouri were utilized as subjects. They were drawn from the medical and surgical wards, and ranged in age from 21 through 60 with a mean age of 37.42 years. Neurological and amputee patients were not included in the sample.

Materials. The rationale and construction of the Pain Apperception Test was presented in a previous article by the writer [9]. Briefly, this test contains 25 pictures which fall into three major divisions, the first of which includes pictures 1 through 9 depicting different situations of felt sensation. The second group contains eight pictures (four pairs) and deals with Anticipation *vs*

---

[1]This paper is based upon a doctoral dissertation submitted to the Department of Psychology of Washington University, St. Louis, Mo.

[2]Grateful acknowledgment is given William Howard French, St. Louis artist, for his creative and painstaking execution of the drawings for the Pain Apperception Test. The writer is indebted to Dr. Ivan N. Mensh, Dissertation Committee Chairman, for his generous help and guidance in the conduct and successful completion of this project; and to Dr. Charles A. Dailey, former Chief Psychologist at St. Louis VA Hospital, whose advice and encouragement gave impetus to the writer's ideas and formulations. My thanks also to Drs. James M. Vanderplas, Carl E. Sherrick, Marion E. Bunch, Rex M. Collier, Walter L. Wilkins, Arnold H. Hilden, Forest C. Orr, and Mr. Albert E. Wibbenmeyer for their constructive criticism of the experimental investigation and its literary presentation.

Felt-Sensation. Counterpart pictures are used here, with one picture of each pair depicting the imminent inflicting of a particular pain (Anticipation) while the other picture depicts the same situation except that the anticipated painful action is being received (Felt-Sensation). Since the counterpart pictures differ only as to whether the pain is imminent or has been inflicted, differences in evaluation of any contrasting pair can be attributed only to the manner in which they differ, the variable of time and proximity (anticipation). The last eight pictures (four pairs) deal with the pain sensation origin (Self-Inflicted *vs.* Other-Inflicted), with one member of each pair depicting a man who has inflicted pain upon himself while the counterpart pictures are identical except that someone else has inflicted the pain upon the pictured person. Differences in evaluation of these counterparts may therefore be related to the manner in which they differ, i.e., whether the pain was self- or other-inflicted. The letters appended to some of the picture numbers, "A" (for "anticipation") and "O" (for "other"), permit easy identification of these pictures and their counterparts.

The pictures include these situations: *Felt Sensation:* 1, bumped head; 2, cut forearm; 3, hit shin; 4, stomachache; 5, gunshot in shoulder; 6, electric shock; 7, headache; 8, stubbed toe; 9, broken leg. *Anticipation vs. Felt-Sensation:* 10-A and 10, fall from ladder; 11-A and 11, ice falling on foot; 12-A and 12, getting hypodermic; 13-A and 13, getting tooth drilled. *Self-Inflicted vs. Other-Inflicted:* 14 and 14-0, thumb mashed; 15 and 15-0, hand burned; 16 and 16-0, finger pricked; 17 and 17-0, removing foreign substance from eye. These pain situations were selected on the basis of a survey of painfulness concepts[10]. Approximately the size of TAT cards, each picture, made by photographing the original lifelike sketches, is 6 by 7½ inches and was mounted on 8½ by 11 inch cardboard backing. Test directions foster the projective frame of reference, and apperceptive response to each of the pictured pain situations was obtained for the qualities of Intensity and Duration by means of two questions for which the subject indicated his multiple-choice answer on a seven-point scale.

Neuroticism was assessed by the Medical Questionnaire (MQ) developed by Eysenck[4] from item analyses of various types of questionnaires having psychological as well as medical importance in the diagnosis of psychoneurosis. The present form of the test is composed of 40 items with the number of "yes" answers constituting the score of the subject. Manifest anxiety was assessed from response to the Taylor Manifest Anxiety Scale[12]. A short pain experience questionnaire was specifically designed for this study, and partly adapted from a similar questionnaire used by Geers[5]. Its principal purpose was to obtain a subjective estimate of amount of pain previously suffered and comparative ability to withstand pain.

Design. Order of presentation of the Pain Apperception Test (PAT) pictures was: *Group A*—1, 2, 3, 4, 5, 6, 7, 8, 9, 10-A, 11, 12-A, 13, 14, 15-0, 16, 17-0; *Group B*—1, 2, 3, 4, 5, 6, 7, 8, 9, 11-A, 10, 13-A, 12, 15, 14-0, 17, 16-0. Subjects were alternately assigned to Group A or Group B to minimize through random selection those variables which might be significant in producing differential response to the Anticipation pictures and their Felt-Sensation counterparts, and to the Self-Inflicted pictures and their Other-Inflicted counterparts. The first nine pictures (felt sensation) were exactly the same for both groups and provided a basis for comparison of any subject(s) with any other subject(s).

The responses to the Intensity and Duration questions for each of the PAT pictures were given a value from 1 through 7. The seven possible Intensity responses were empirically established by use of Edwards' method of successive intervals[3] and the seven multiple-choice Duration responses representing a logical temporal continuum. When comparison of individual pictures was concerned, each subject's score could vary from 1 through 7 with the means reflecting average scores for a particular picture. When scores for the first nine pictures were taken to represent response to the PAT, these scores were obtained for each subject by summing his responses, for Intensity or for Duration, on each of these nine pictures. Thus the 1-9 score for any subject may range from 9 through 63 for Intensity or for Duration, and mean scores represent the average for a particular group on these pictures.

The standard order of test administration and approximate average time spent on each was: Pain Apperception Test, 15 to 20 minutes; Medical Questionnaire, 20 minutes; Biographical Inventory (Taylor MAS), 50 minutes; pain experience questionnaire, 2 minutes. In the analysis of results, statistical techniques were employed as advocated by Edwards[2] and by McNemar[8]. Homogeneity of variance was taken into account whenever necessary and, for the data of all hypotheses, the observed score distributions were subjected to Chi-square goodness of fit test to evaluate the approximation of data to the theoretical normal curve.

Results

Reliability. As shown in Table 1, split-half reliability of the PAT was found to vary from .56 to .85 depending upon the experimental group and its size, the pain dimension, and the number of pictures included in the calculation by Spearman-Brown formula. These reliability coefficients are sufficiently high to denote consistency of response to the PAT, and to suggest reliability of results of this study. Split-half reliability coefficients obtained for the criterion tests of neuroticism (Medical Questionnaire) and manifest anxiety (Taylor MAS) were quite high (.88 and .93, respectively), and compare favorably with results of earlier studies.

TABLE 1. RANGE AND SPLIT-HALF RELIABILITY OF PAIN APPERCEPTION TEST SCORES[a]

PAT Scoring Category	Subject Sample	Response Range	Split-half Reliab.
Intensity	Group A	49-83	.79
	Group B	49-79	.67
	All Ss	27-48	.56
Duration	Group A	47-89	.85
	Group B	45-83	.84
	All Ss	27-51	.66

[a]For Groups A and B, N = 50 each for the 17 pictures administered each group; for All Subjects, N = 100 for pictures 1 through 9.

Neuroticism and Manifest Anxiety. PAT scores showed a significant positive relationship to MQ scores and to Taylor MAS scores (Table 2). Product-moment correlations ranged from .31 to .43 and were significant at the .01 and .001 confidence

TABLE 2. PRODUCT-MOMENT AND PARTIAL CORRELATIONS AMONG PAIN APPERCEPTION TEST SCORES,[a] MEDICAL QUESTIONNAIRE SCORES, AND TAYLOR MAS SCORES (N = 100)

Product-moment *r*		Partial *r*	
r MQ, Int	= .43***	*r* MQ, Int (Dur)	= .34***
r MQ, Dur	= .31**	*r* MQ, Dur (Int)	= .24*
r MAS, Int	= .39***	*r* MAS, Int (Dur)	= .29**
r MAS, Dur	= .32**	*r* MAS, Dur (Int)	= .18

[a]PAT scores are for total response to pictures 1 thru 9, for Intensity or for Duration.

*Significant at .05 level.
**Significant at .01 level.
***Significant at .001 level.

levels. The partial r's usually do not fall below the .05 confidence level, and indications are that each of the variables of Intensity and Duration contribute a greater proportion to their own scores than does the influence of the other of these two variables. Interaction effects between Intensity and Duration are indicated and this is in accord with expectation based on clinical observation. It is also evident that prediction involving pain apperception is enhanced if the interaction between these two variables is not removed from correlation coefficients.

Considering the types of tests involved, the reliability of obtained responses, and the correlations resulting from comparison of their scores, consistency in individual response as to the Intensity and Duration of apperceptive judgment of pain situations and the variables of neuroticism and manifest anxiety is strongly indicated by these results.

Anticipation and Felt-Sensation. Results presented in Tables 3, 4, and 5 permit the individual consideration of each counterpart pair, since each represents a different stimulus value. Important factors to consider in the first pair of pictures include: (a) the impending pain of picture 10-A is reasonably certain of being felt (picture 10) in a very short time, i.e., the time required for the man to fall one or two feet; (b) the painfulness connoted by this situation was judged to be fairly high by the total subject sample; and (c) the pictured man is apparently helpless to prevent his accidental injury. PAT scores reveal a tendency for the total sample to evaluate the Anticipation as more painful in Intensity and Duration than the Felt-Sensation (Table 3); the significantly greater painfulness evaluation of Anticipation by those designated as neurotic (Table 4); and the tendency for Anticipation to be evaluated as more painful by subjects with high manifest anxiety (Table 5).

TABLE 3. MEAN PAIN APPERCEPTION TEST SCORES FOR ANTICIPATION AND FELT-SENSATION PICTURES: ALL SUBJECTS (N = 50 PER SCORING CATEGORY)

PAT Scoring Category	Anticipation Pict. No.	M	SD	Felt-Sensation Pict. No.	M	SD	t
Intensity	10-A	4.92	1.55	10	4.40	1.92	1.49
Duration	10-A	5.36	1.67	10	4.98	1.50	1.19
Intensity	11-A	3.80	1.65	11	4.96	1.60	3.57**
Duration	11-A	3.74	1.61	11	4.72	1.20	3.45*
Intensity	12-A	2.04	.70	12	2.16	.74	.83
Duration	12-A	2.00	.73	12	2.00	.76	0.00
Intensity	13-A	2.64	1.12	13	2.88	1.39	.95
Duration	13-A	2.48	.93	13	2.62	.94	.74

*Significant at .01 level.
**Significant at .001 level.

Evaluation of the second counterpart pair should take cognizance of these aspects: (a) the impending hurt in picture 11-A should occur very shortly, i.e., the time required for the ice to fall about six or eight inches; (b) the Intensity and Duration of the pain were judged by the total sample to be moderately high for these pictures; and (c) the injury would be, and is (picture 11), accidental. Results for this counterpart pair showed the Felt-Sensation scores for Intensity and Duration to be significantly greater for the total sample (Table 3), for those designated as neurotic (Table 4), and for those subjects with high manifest anxiety (Table 5). A possible explanation of these results may be found in a closer evalutation of the pictures, and comparison with the preceding pair. Unlike picture 10-A (man falling), picture 11-A is not necessarily of a man helpless to alter his immediate fate. Accidental injury is unwelcome and the ice man could jerk his foot entirely or partly from the path of the falling ice. Also, though the picture objectively suggests otherwise, defensive perception (apperception) could make it appear that the falling ice is destined to miss the unmoving foot, or strike it tangentially.

The third pair of counterpart pictures has these aspects: (a) the temporal imminence of picture 12-A is not so immediate, i.e., the time required to insert the hypodermic needle; (b) the Intensity and Duration of the pain were evaluated as fairly low by the total subject sample; and (c) the intended effect of the painful action is beneficial rather than harmful. Results indicate no significant differences in

TABLE 4. MEAN PAIN APPERCEPTION TEST SCORES FOR ANTICIPATION AND FELT-SENSATION PICTURES: NEUROTIC SUBJECTS[a]

PAT Scoring Category	Anticipation Pict. No.	N	M	SD	Felt-Sensation Pict. No.	N	M	SD	t
Intensity	10-A	12	5.58	.79	10	14	4.57	1.74	1.95*
Duration	10-A	12	6.08	1.00	10	14	5.21	1.19	2.02*
Intensity	11-A	14	4.14	1.61	11	12	5.25	.75	2.30*
Duration	11-A	14	4.07	1.38	11	12	5.00	1.28	1.75*
Intensity	12-A	12	2.08	.52	12	14	2.14	.86	.22
Duration	12-A	12	2.00	.60	12	14	2.14	1.10	.40
Intensity	13-A	14	2.93	1.00	13	12	3.42	1.44	1.02
Duration	13-A	14	3.00	.88	13	12	3.17	.84	.50

[a]The approximate upper quartile (N = 26) of Medical Questionnaire scores was labeled "neurotic".
*Significant at .05 level.

TABLE 5. MEAN PAIN APPERCEPTION TEST SCORES FOR ANTICIPATION AND FELT-SENSATION PICTURES: HIGH ANXIETY SUBJECTS[a]

PAT	Anticipation				Felt-Sensation				
Scoring Category	Pict. No.	N	M	SD	Pict. No.	N	M	SD	*t*
Intensity	10 A	16	5.43	.96	10	12	5.00	1.35	.98
Duration	10-A	16	5.81	1.11	10	12	5.08	1.16	1.69
Intensity	11-A	12	4.42	1.16	11	16	5.44	.73	2.85**
Duration	11-A	12	4.25	.97	11	16	5.12	1.02	2.28*
Intensity	12-A	16	2.19	.54	12	12	2.42	.79	.92
Duration	12-A	16	2.06	.57	12	12	2.50	1.00	1.47
Intensity	13-A	12	2.92	1.08	13	16	3.56	1.67	1.16
Duration	13-A	12	2.92	1.00	13	16	3.12	.81	.59

[a]The approximate upper quartile (N = 28) of Taylor MAS scores was labeled "high manifest anxiety".

*Significant at .05 level.

**Significant at .01 level.

scores for these two pictures for the total sample (Table 3), for neurotics (Table 4), or for those with high manifest anxiety (Table 5). It may be that the utilized sample of hospitalized male veterans is comparatively inured to and unthreatened by "shots," the comparatively low intensity and duration of such pain does not elicit differential apperception between Anticipation and Felt-Sensation, and/or the beneficial intent of the necessary pain of injection minimizes the threat component.

The fourth pair of counterpart Anticipation and Felt-Sensation pictures has these characteristics: (a) temporal imminence of picture 13-A is not so immediate, i.e., time until the dental drill is inserted into the mouth and operated; (b) the total subject sample evaluated the painfulness of this situation as somewhat low; and (c) the painful action has a beneficial intent. As with the preceding counterpart pair, no significant differences were found for the three comparisons made. A possible explanation of these results may lie in the beneficial intention causing the pain, and the factor of control may also be important since the "patient" could have an analgesic or indicate his discomfort to the dentist.

Of the four counterpart situations, the fall from the ladder (pictures 10-A and 10) most clearly yields greater PAT scores for the Anticipation as compared to the Felt-Sensation; and it is this situation which is characterized by a great amount of physical pain, high imminence, lack of control, and an absence of beneficial intent. PAT scores for pictures 11-A and 11 (ice falling on foot) were consistently and significantly greater for the Felt-Sensation. This pain situation is also characterized by much physical painfulness (if and when the ice hits the foot), high imminence, and the absence of beneficial intent. However, the element of control does seem to be present in the Anticipation picture and this may account for the lower scores obtained in comparison to the Felt-Sensation where considerable hurt has been sustained. PAT scores for pictures 12-A and 12 (hypodermic) and 13-A and 13 (tooth drilling) did not reveal significant differences for the counterpart comparisons. Both sets of pictures are apparently characterized by comparatively low physical painfulness and a lesser degree of imminence than the other two pairs, the control factor is at least partially present, and the underlying intent is beneficial. It is possible that, under these circumstances, differences in reaction to the anticipation and to the culmination of a painful act are comparatively slight.

The PAT scores obtained for the four counterpart pairs suggest that the following variables are important to understanding anticipatory pain reactions: (a) the underlying physical painfulness usually associated with the act; thus, most people would agree that a smashed toe hurts more than a pin prick; (b) the proximity of the pain; the more imminent the pain, the greater the emotional trauma associatedwith

it; (c) the control factor—pain is more bearable when its inflicting, duration, and/or cessation are capable of modification by the sufferer; and (d) the underlying intention; thus, inflicting pain for a beneficial purpose might be expected to be less traumatic to the individual than a similar inflicting of pain resulting from accident or evil intent. It also seems evident from the results shown in Tables 4 and 5 that neuroticism and manifest anxiety are important in the apperception of anticipatory pain situations; and a general increase in apperceived pain as a function of neuroticism and/or manifest anxiety was indicated by the significant positive correlations found between these variables (Table 2). To the above four factors (physical pain, imminence, control, and intent) which are oriented toward the interpersonal and physical situation, it therefore is reasonable to add the fifth and more intrapersonal variable of neuroticism and/or manifest anxiety. That there is a dynamic interrelationship between all these factors is very likely.

Self-Inflicted and Other-Inflicted. Results were analyzed for these counterpart pairs in the same way as for the Anticipation vs. Felt-Sensation counterparts, and are available in Tables A, B, and C.[3] Significant differences were found in the evaluation of pictures 14 and 14-0 but the independent drawing (no overlay utilized) of the two pictures in this instance did not assure adequate stimulus control and made these results inconclusive.

Pictures 15 and 15-0 represent a moderately painful situation as indicated by the mean scores of the entire sample population, and both pictures depict accidental injury (hand burned by hot liquid). Results showed higher PAT scores for the Other-Inflicted picture in the case of all subjects, neurotic subjects, and high manifest anxiety subjects; and these differences were statistically significant by t test, at the .05 or .01 level, for three of the six group comparisons. A possible explanation for this difference may be found in the social implications. The burned hand injury of picture 15-0 was perpetrated by a presumably unknown person, the waiter, whose skill should not permit such accidents to occur. Aggression toward the errant waiter, and magnification of the sustained injury, might thus ensue. The apperception of lesser pain when the burn is Self-Inflicted (picture 15) may signify a conscious or unconscious diminution of painful effects when the person is responsible for his own pain, and perhaps stupidly so, in a public social setting. The control factor may also be important since the pictured man has greater control of the situation when he is burning himself, as compared to when the other person is causing him to be burned.

No significant differences were found between pictures 16 and 16-0, and counterpart comparison is complicated by the motivation underlying each. Picture 16 shows an accidental needle prick by the self while picture 16-0 shows a deliberate pricking of the finger by a professional person (blood sample) whose intention is to benefit the recipient.

Pictures 17 and 17-0 represent relatively mild painfulness and depict a painful but beneficial action, i.e., removing a foreign body from the eye. No significant differences in PAT scores were found for the total sample, for neurotics, or for subjects with high manifest anxiety.

Subjective Past Pain Evaluation. Subjective estimates regarding comparative pain suffered during one's lifetime, and regarding comparative ability to stand pain were obtained from questions 4 and 5 of the pain experience questionnaire. The three possible responses to each question permitted the subject a choice of above average, average, or below average. Results are available in Tables D and E[3]. Variance analyses did not disclose significant differences between total PAT scores for the first nine pictures when group comparisons were made concerning the comparative amount of previous pain experience (Table D) and concerning comparative

[3]Three 1-page tables giving data for the entire sample, for Ss designated neurotic, and for Ss with high manifest anxiety, and two 1-page tables giving data for past pain experienced and for comparative ability to stand pain have been deposited with the American Documentation Institute. Order Document No. 5544, remitting $1.25 for each 35-mm. microfilm or $1.25 for 6 by 8 in. photocopies.

ability to stand pain (Table E) as indicated by the subjects. However, mean PAT scores for the different subject groups did reveal a definite tendency to apperceive greater pain in the pictures as the subjective amount of past pain suffered increased, or as subjective ability to stand pain decreased.

Discussion

In accord with its primary purpose, results of this study have indicated intra-individual consistency in the apperception of pain in pictures, self-report of manifest anxiety, self-report of physical and psychological complaints usually associated with neuroticism, and, to some extent, subjective opinion regarding past pain experience and ability to stand pain. The distributions of group scores for the many relationships studies were found to be normal by statistical test. Results obtained in this study may therefore be generalized to similar male patient populations. The Pain Apperception Test is suitable for experimentation with other populations, and for the investigation of pain apperception as related to such variables as psychiatric diagnosis, personality attributes, age, sex, race, intelligence, occupation, physical diagnosis, cultural background, socio-economic status, and others, as well as those variables considered in the present experiment.

Picture Changes. Results of this study have indicated that certain of the Self- vs. Other-Inflicted counterpart pictures of the PAT were not sufficiently definitive as to stimulus characteristics. Accordingly, pictures 14-0, 16, and 16-0 were replaced. Pictures 5 and 14 were interchanged; the new Other-Inflicted counterpart (picture 14-0) is exactly the same as picture 14 except the rifle is held by another person whose hands only are visible. A new picture was also drawn by the original artist to replace pictures 16 and 16-0 (pricked finger); the replacements depict two men fly-fishing and these counterparts differ only as to whether the man in pain has been hooked in the neck by his own line (Self-Inflicted) or by the line of the other man (Other-Inflicted). It is felt that these changes will introduce more control and enhance the discriminatory power of the Pain Apperception Test.

Summary

This research studied the efficacy of the Pain Apperception Test as a research tool by investigating the consistency of individual response as to pain apperception, neuroticism, manifest anxiety, and self-appraisal. The research design involved the PAT, a medical questionnaire, Taylor Manifest Anxiety Scale, and a pain experience questionnaire. Data from responses to these instruments permitted studying the relationship of pain apperception to personality characteristics, and to subjects' ratings of their comparative past pain experience. Relationships also were examined regarding the comparative apperception of pain in the counterpart PAT pictures, thus investigating Anticipation vs. Felt-Sensation, and Self-Inflicted vs. Other-Inflicted pain situations. One hundred hospitalized, male patients, 21 through 60 years of age, comprised the subjects. Results and conclusions were:

1. Split-half reliability of the essentially projective Pain Apperception Test was found to range from .56 to .85, depending upon the number of subjects and pictures, and the pain dimension.

2. There was a significant positive correlation between the apperception of pain in pictures and neuroticism as determined from response to the Medical Questionnaire.

3. There was a significant positive correlation between the apperception of pain in pictures and manifest anxiety as determined from response to the Taylor Manifest Anxiety Scale.

4. No significant relationships were found between PAT scores and the opinion of subjects regarding the comparative amount of pain suffered in their lifetimes, or regarding their comparative ability to stand pain. A definite tend-

ency to apperceive greater pain in the pictures as subjective amount of past pain suffered increased, or as subjective ability to stand pain decreased, was indicated.

5. Differential PAT scores for the counterpart Anticipation and Felt-Sensation pictures suggest the following factors are important in assessing reactions to anticipated pain: (a) the degree of physical painfulness commonly associated with the impending injury; (b) the proximity or imminence of the pain; (c) the degree to which the intended recipient can control the receiving, duration, and/or termination of the impending pain; (d) the intention or effect (beneficial or harmful) of the impending pain; and (e) the degree of neuroticism and/or manifest anxiety possessed by the recipient of the painful action. It is felt that these factors all contribute to the particular pattern of past associations, present personality, and situational variables responsible for a given individual's reaction to a specific situation wherein pain is anticipated.

6. Differential response to the counterpart Self-Inflicted and Other-Inflicted pictures was not very definitive, but did indicate that social implications may be important in determining reactions to painful stimuli. Recent changes in this group of pictures may yield more definite conclusions.

These results indicate intra-individual consistency in apperceptive evaluation of the pictured pain situations of the PAT, self-report of physical and psychological complaints associated with neuroticism, self-report of manifest anxiety, and, to some extent, subjective opinion of past pain suffered and ability to stand pain. Thus, the apperception of pain in pictures is evidently a reaction consistent with important personality attributes, rather than a unique experience in the psychology of the individual. The Pain Apperception Test is evidently suited to research involving the psychological dimensions of pain, and the use of this test with various populations and problems should increase our understanding of pain and its psychological aspects.

References

1. Chapman, W. P. Measurement of pain sensitivity in normal control subjects and in psychoneurotic patients. *Psychosom. Med.*, 1944, *6*, 252-257.
2. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
3. Edwards, A. L. The scaling of stimuli by the method of successive intervals. *J. appl. Psychol.*, 1952, *36*, 118-122.
4. Eysenck, H. J. *Dimensions of personality.* London: Rutledge Kegan Paul Ltd., 1948.
5. Geers, J. B. Pain thresholds in human subjects. Unpublished doctoral dissertation, Univer. of Missouri, 1954.
6. Hill, H. E., Kornetsky, C. H., Flanary, H. G. and Wikler, A. Studies on anxiety associated with anticipation of pain. I. Effects of morphine. *Arch. Neurol. Psychiat.*, 1952, *67*, 612-619.
7. Kornetsky, C. H. Effects of anxiety and morphine on the anticipation and perception of painful radiant stimuli. *J. comp. physiol. Psychol.*, 1954, *47*, 130-132.
8. McNemar, Q. *Psychological statistics.* (2nd Ed.) New York: Wiley, 1955.
9. Petrovich, D. V. The Pain Apperception Test: a preliminary report. *J. Psychol.*, 1957, *44*, 339-346.
10. Petrovich, D. V. A survey of painfulness concepts. *J. clin. Psychol.*, 1958, *14*, 288-291.
11. Schilling, R. F. and Musser, M. J. Pain reaction threshold in psychoneurotic patients. *Amer. J. med. Sci.*, 1948, *215*, 195-197.
12. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, *48*, 285-290.

GRAPHIC REPRESENTATION OF EMOTIONAL FEELINGS[1]

GEORGE A. PETERS
Psychological Research Associates, Inc.
Encino, California

PHILIP R. MERRIFIELD
University of Southern California
Los Angeles, California

PROBLEM

In personality appraisal and the differential diagnosis of psychopathology, there has been widespread use of projective tests requiring the production of written or graphic material. Interpretations regarding the emotional feelings of the respondant are often based upon certain characteristics of the lines which are drawn, such as the pressure of the stroke, the angularity of the line, and tendencies to deviate from the horizontal plane. Assumptions of this nature are implicit in such testing devices as the Bender Visual Motor Gestalt Test (1, 4, pp. 331-334), the Draw-A-Person type tests (6, pp. 281-282), finger painting tests (7, p. 406; 5, pp. 418-419), and graphology (8). The validity of such interpretive assumptions rests upon theoretic formulations or deductions and empirical case study evidence from pathological populations. This report presents evidence from a "normal" or non-pathological sample regarding the ways in which certain emotional feelings were graphically represented.

PROCEDURE

A group of 24 adjectives (2) were presented to a sample consisting of 188 male United States Air Force air crew trainees who were instructed to draw a single line best expressing the meaning of each given word (3). The adjectives were printed on two pages with space provided after each word for the line drawing. It was administered as a group test with ample time for each examinee to complete all items. The mean age of the sample was 20.75 (S. D. 1.54 yrs.), and the mean number of years of education completed was 13.32 (S. D. 1.27 yrs.). The test reliability was 0.63 (unmatched split-half with Spearman-Brown correction for length).

The responses of each examinee were subsequently categorized (2), relative to his other responses, on three dimensions; form (straight, curvilinear, wavy, or angular), direction (horizontal, upward, or downward), and intensity of line (light, medium, or heavy stroke.) To classify the responses obtained for each word, a weight was assigned in accordance with the obtained frequency in a scoring category when compared to the mean and standard deviation of the total distribution for all words. For example, if a word received a weighted score of 4 for horizontal direction, it means that the frequency of horizontal responses to that word was more than that found at two sigmas above the mean of the horizontal responses for all words combined. A weighted score of 3 indicates that the frequency of responses in a particular scoring category for a given word was within the first standard deviation above the mean of responses for all words (see tables 1 and 2).

RESULTS

Based upon the frequency of responses in each of the scoring categories (see Tables 1 and 2), it was possible to make certain inferences or hypotheses regarding the general meanings apparently associated with the way in which simple lines are drawn. These hypotheses were formulated to aid in the design of future studies which may substantiate or extend these findings.

The *direction* of the line seems related to the *social desirability* of the behavior or feeling represented by the word. A downward tendency is apparently used more frequently to describe the less socially desirable attributes associated with the

[1]This study was based upon the data collected under Contract N6ONR-23810, Project 150-044, between the Personnel and Training Branch of the Office of Naval Research and the University of Southern California. J. P. Guilford, responsible investigator and director of the Project on the Aptitudes of High-level Personnel, originated the test instrument used in this study. The ideas expressed in this paper are our own and do not necessarily represent the opinions of the Office of Naval Research.

TABLE 1. LINE FORM AND DIRECTION FOR GIVEN WORDS

Categories	Horizontal	Upward	Downward
Angular	Lively (4, 3) Angry (4, 2) Agitating (3, 3)	Fiery (4, 3) Furious (4, 3) Angry (4, 2)	Cruel (3, 3) Angry (4, 2)
Wavy	Playful (4, 3) Jolly (4, 3) Delicate (4, 3) Agitating (4, 3) Lively (3, 3)	Joyous (3, 3)	
Curvilinear	Tranquil (3, 3) Gentle (3, 3)	Joyous (3, 3)	Sad (4, 4) Faint (3, 4) Sorrowful (3, 4) Lazy (3, 3)
Straight	Quiet (4, 4) Grave (3, 3) Idle (3, 3)	Forceful (4, 4) Earnest (3, 4) Hard (3, 3)	Dead (4, 3)

The numbers following each word indicate the standard score value for form and direction, respectively, on a scale of 1 to 4. The highest value, 4, indicates the greatest frequency of response.

words cruel, lazy, sad or sorrowful, and dead. Conversely, an upward tendency is used more frequently to describe the more socially desirable attributes associated with the words earnest, joyous, and strong.

The *form* of the line is apparently related to the degree of *overt activity* suggested by the word. The more angular or irregular lines are used to represent words such as lively, agitating, angry, and furious, whereas a straight line is used to represent the words grave, idle, quiet, and hard.

The *intensity* of the line is apparently related to the *intensity, persistence,* or *strength* of the emotional feeling represented by the word. Heavy lines were used to represent words such as angry, furious, and cruel; in contrast, a light line was used to represent the words gentle, delicate, and quiet. Words such as forceful, hard, and

TABLE 2. LINE INTENSITY FOR GIVEN WORDS

Heavy	Medium	Light
Cruel, 4	Playful, 4	Quiet, 4
Forceful, 4	Lively, 4	Gentle, 4
Furious, 4	Jolly, 4	Faint, 4
Strong, 4	Joyous, 4	Delicate, 4
Fiery, 3	Agitating, 3	Tranquil, 3
Hard, 3	Tranquil, 3	Idle, 3
Angry, 3	Sorrowful, 3	
	Sad, 3	

The numbers following each word indicate the standard score value on a scale of 1 to 4 with the highest value indicating the greatest frequency of response.

strong, although not suggestive of a display of overt activity (they were represented by straight lines), nevertheless do possess a quality of strength, intensity, or persistence and are, therefore, represented by heavy lines. Similarly, the words lively and agitating are suggestive of overt activity (angular lines) but only moderate strength, intensity, or persistence (lines of medium intensity).

To gain further evidence regarding the validity of these hypotheses, a number of stimulus words were paired for purposes of comparison. Chi Square tests of the null hypothesis among the selected stimulus word pairs revealed some markedly significant differences which tend to support the tentative hypotheses regarding the graphic representation of emotional feelings (see Table 3).

TABLE 3. PAIRED STIMULI TESTS OF HYPOTHESES

Stimulus Pair	Direction of Line Hyp: *Social Desirability*				Form of the Line Hyp: *Overt Activity*				Intensity of Line Hyp: *Strength*			
	Downward	Horizontal	Upward	Sign	Angularity	Curvilinear*	Straight	Sign	Heavy	Medium	Light	Sign
Sorrowful	123**	60	5	$X^2=182.70$	6	161	21	$X^2=27.94$	23	110	55	$X^2=23.98$
Vs.				$P = <.001$				$P = <.001$				$P = > .001$
Joyous	6	103	79	$\phi = .70$	32	151	5	$\phi = .27$	36	134	18	$\phi = .25$
Fiery	11	99	78	$X^2=11.74$	87	81	20	$X^2=81.76$	130	56	2	$X^2=12.80$
Vs.				$P = <.01$				$P = <.001$				$P = <.01$
Hard	30	100	58	$\phi = .18$	47	39	102	$\phi = .47$	159	28	1	$\phi = .18$
Strong	7	83	98	$X^2=36.00$	44	70	74	$X^2=49.10$	180	8	0	$X^2=335.52$
Vs.				$P = <.001$				$P = <.001$				$P = <.001$
Gentle	13	132	43	$\phi = .31$	51	130	51	$\phi = .36$	3	71	114	$\phi = .94$
Cruel	63	90	35	$X^2=62.06$	85	74	29	$X^2=83.86$	168	19	1	$X^2=1.14$
Vs.				$P = <.001$				$P = <.001$				$P = > .10$
Forceful	9	93	86	$\phi = .41$	48	26	114	$\phi = .47$	171	17	0	$\phi = .06$
Lively	6	140	42	$X^2=14.78$	0	95	0	$X^2=216.26$	34	147	7	$X^2=169.54$
Vs.				$P = <.001$				$P = <.001$				$P = <.001$
Quiet	14	157	17	$\phi = .20$	121	66	121	$\phi = .76$	9	51	128	$\phi = .67$
Angry	37	100	51	$X^2=8.96$	98	62	28	$X^2=31.02$	159	29	0	$X^2=57.90$
Vs.				$P = <.02$				$P = <.001$				$P = <.001$
Agitating	20	126	42	$\phi = .15$	64	114	10	$\phi = .29$	87	97	4	$\phi = .39$
Lazy	79	92	17	$X^2=96.48$	9	147	32	$X^2=76.10$	16	120	52	$X^2=79.64$
Vs.				$P = <.001$				$P = <.001$				$P = <.001$
Earnest	12	89	87	$\phi = .51$	27	63	98	$\phi = .45$	90	90	8	$\phi = .46$
Faint	92	66	30	$X^2=60.14$	4	133	51	$X^2=129.98$	5	50	133	$X^2=310.40$
Vs.				$P = <.001$				$P = <.001$				$P = <.001$
Furious	23	106	59	$\phi = .40$	104	62	22	$\phi = .59$	173	15	0	$\phi = .91$

*Includes both Curvilinear and Wavy
**Obtained cell frequences

Discussion

Although these results apparently indicate that there is some common agreement as to how emotional feelings can be expressed in line drawings, this is probably to be expected in view of the communication values inherent in artistic productions. The hypotheses derived from this empirical data may tend to clarify or make explicit some of the techniques utilized for such artistic communication. The data also suggest that there is a fairly general understanding of such communicative devices. There is, however, a logical fallacy in then assuming that individual drawings can be used clinically for the identification of emotional feelings. That is, it does not imply that emotional feelings will actually modify an unsuspecting individual's drawings in such a manner as to permit identification of his emotional feelings. For example, most people will draw a downward line to represent the word sad. But, under some other instructions, it cannot be inferred that individuals who draw downward lines do so because they are sad. This logical fallacy suggests why such interpretations are so commonly formulated even without adequate substantiating data.

There are a number of very interesting and theoretically important implications in the assumption that an individual's emotional feelings will reliably modify his graphic representations in a predictable manner. But it is essential that this assumption be adequately supported by appropriate scientific investigation of (a) possible changes in an individual's drawings under experimentally induced or stress evoked alterations in his emotional feelings, (b) comparative differences in normative data between groups categorized as average or extreme in the symptomatic expression of the emotional feelings, (c) ipsative data of temporal variations in graphic productions associated with naturally occurring fluctuations in an individual's emotional tone, and (d) the influence of such things as sex differences, cultural factors, and intellectual ability. If such scientific evidence seems promising, a greater number of hypotheses or significant predictive indices could be derived as an aid in the clinical appraisal of emotional feelings and for use in the development of more refined diagnostic instruments.

Conclusions

The results of this study indicate that there is some general agreement in a nonpathological population regarding the graphic representation of emotional feelings. Additional research is indicated to further investigate the hypothesis than an individual's emotional feelings will, reliably, modify his graphic productions in a predictable manner.

References

1. Bender, Lauretta. A visual motor gestalt test and its clinical use. *Res. Monogr. No. 3.* New York: American Orthopsychiat. Assoc. 1946.
2. Guilford, J. P. and Guilford, Ruth B. A prognostic test for students in design. *J. appl. Psychol.*, 1931, *15*, 335-345.
3. Guilford, J. P., Berger, R. M. and Christensen, P. R. A factor-analytic study of planning, II. Administration of tests and analysis of results. *Rep. psychol. Lab.*, No. 12. Los Angeles: Univer. of Southern California, May 1955.
4. Halpern, Florence. The Bender Visual Motor Gestalt Test. In Anderson, H. H. and Anderson, Gladys L. (Eds.) *An Introduction to projective techniques.* New York: Prentice-Hall, 1951, 324-340.
5. Kadis, A. Finger-painting as a projective technique. In Abt, L. E. and Bellak, L. (Eds.) *Projective psychology.* New York: Knopf, 1952, 403-431.
6. Levy, S. Figure drawing as a projective test. In Abt., L. E. and Bellak, L. (Eds.) *Projective psychology.* New York: Knopf, 1952, 257-297.
7. Napoli, P. Finger Painting. In Anderson, H. H. and Anderson, Gladys (Eds.) *An introduction to projective techniques.* New York: Prentice-Hall, 1951, 386-415.
8. Wolfson, Rose. Graphology. In Anderson, H. H. and Anderson, Gladys L. (Eds.) *An introduction to projective techniques.* New York: Prentice-Hall, 1951, 416-456.

THE VALIDITY OF THE EDWARDS PERSONAL PREFERENCE SCHEDULE IN THE MEASUREMENT OF DEPENDENCY-REBELLIOUSNESS

MARVIN ZUCKERMAN

Institute of Psychiatric Research, Indiana University Medical Center, Indianapolis

The Edwards Personal Preference Schedule (PPS) [2] attempts to eliminate the factor of social desirability response set (the tendency of a testee to answer an inventory according to the perceived social desirability of the items) which seems to have a marked influence on most previous inventories. The scales in this inventory are a sample of needs from the Murray system [3]. Several of the PPS variables seem to bear on a fundamental concept in clinical work, "Dependency". Deference, Succorance, and Abasement are characteristics of a "Dependent Personality". As the opposite extreme of Dependency we might expect to find "Rebelliousness", perhaps expressed in needs like Autonomy, Dominance, and Aggression. Bernardin and Jessor [1] used the Deference and Autonomy scales of the PPS as criteria for defining groups of "Dependent" and "Independent" subjects. The Dependents were found to differ significantly from Independents on two experimental tasks: Dependents made more errors on a maze task under conditions of negative reinforcement, and asked for help more often in a difficult problem solving situation. In a previous study by Zuckerman and Grosz [4], "suggestible" subjects (defined by response to the Sway Test) were found to be significantly lower on the Autonomy scale of the PPS than "non-suggestible" subjects. There was a tendency (between .10 and .05) for the "suggestible" group to be higher on the Succorance scale of the PPS.

These studies suggest that the PPS might be useful in predicting behavioral tendencies related to the concept of "Dependency". In the course of a study of personality correlates of the Acquiescence Response set [5], the PPS was given to the subjects and they rated each other on certain behavior traits. Four of these traits, "Rebelliousness", "Conformity", "Submissive to Authority", and "Dependency", are relevant to the concept of Dependency-Rebelliousness. The latter three are assumed to exemplify the positive pole of Dependency. The group of subjects identified by their peers as "Rebellious" would be expected to differ from each of the other three criterion groups while little difference would be expected between these three groups. The "Rebellious" group was expected to be higher on the scales, Autonomy, Dominance, and Aggression, and lower on Deference, Succorance, and Abasement. Furthermore, a ratio of one set of combined PPS scales to the other set of combined PPS scales was expected to be effective in separating the "Rebellious" group from the three Dependency groups.

Procedure

Sixty-three sophomore-year student nurses were given the PPS. Eight months later all of these subjects were mailed a form asking them to nominate persons, from the list of names in their group, who best fitted the behavioral categories. The description of each of the four relevant categories read as follows:

1. Which girl on the list would you describe as the most submissive to authority, *i.e.*, one who always does what is asked of her without question or argument?
2. Which girl on the list would you describe as the greatest "conformist", *i.e.*, a conformist as one who goes along with the group on most matters?
3. Which girl on the list would you describe as most "dependent", *i.e.*, one who constantly leans on other people, asking for help and sympathy?
4. Which girl on the list would you describe as the most "rebellious", *i.e.*, one who frequently rebels against orders and refused to comply with what is expected of her?

Two additional categories used were "Prejudiced" and "Hostile", but these were considered less relevant for the concept of Dependency. Some mention will be made of these groups in relation to the Aggression scale.

About two-thirds of the group returned the forms. For each trait a group was formed using the subjects who had been nominated by two or more of their peers as typifying that trait. This criterion was decided on as the minimum cutting point which would exclude idiosyncratic judgments and yet give the minimum N necessary for any sort of comparisons. Ten subjects fell into the Submissive group, seven in the Conformist group, eight in the Dependent group, and seven in the Rebellious group. There was some overlap between the Submissive, Conformist, and Dependent groups, *i.e.*, some subjects were nominated by peers in two or three of these categories. This was expected since all of these traits were conceptualized as aspects of the general trait of Dependency. The fact of overlap precludes any comparisons between these groups. There was no overlap between subjects in the Rebellious group and those in the three Dependency groups. Comparisons were made only between the Rebellious group and each of the three Dependency groups. Although the N's in these groups were small the differences were clear-cut enough to warrant discussion of the findings. Cross-validation of findings would be highly desirable, and it is hoped that some reader will undertake this task. It should be noted that the groups used here probably represent marked extremes since the subjects were asked to nominate the one individual out of 63 most typifying the trait.

Combination and ratio scores from the PPS were formed in the following manner: each subject's raw scores on the six PPS scales were converted to standard scores using the conversion tables in the Edwards manual. The standard scores for the Deference, Succorance, and Abasement scales were added together to form one combination score (DSA). The standard scores for the Autonomy, Dominance, and Aggression scales were added together to form the other combination score (ADA). The "Dependency-Rebelliousness" ratio (D-R) was computed for each subject by dividing his DSA by his ADA score. According to the hypothesis the ratio should be high for Dependent subjects and low for Rebellious subjects.

Results and Discussion

Before discussing the main results, one point of interest lies in a comparison between the total group of 63 student nurses and Edward's standardization group of 749 college girls. Most of the student nurse group had had some college as part of their nursing training program and were about equivalent in age to Edward's group. The comparison between these groups is contained in Table 1. The Student Nurse group was significantly higher than the College group on Abasement and Nurturance and lower on Dominance (in all three p was less than .01). The Student Nurse group

Table 1. Difference Between Total Student Nurse Group and Edward's Female College Group

EPPS Scale	Student Nurses (N = 63)		College (N = 749)		
	M	SD	M	SD	t
Achievement	12.41	3.87	13.08	4.19	.98
Deference	12.42	3.64	12.40	3.72	.03
Orderliness	10.65	3.06	10.24	4.37	.65
Exhibition	13.86	4.08	14.28	3.65	.63
Autonomy	10.97	4.15	12.29	4.34	1.82*
Affiliation	17.90	4.07	17.40	4.07	.72
Intraception	18.13	5.10	17.32	4.70	.96
Succorance	12.63	4.65	12.53	4.42	.13
Dominance	11.67	3.85	14.18	4.60	3.49***
Abasement	17.46	3.66	15.11	4.94	3.20***
Nurturance	18.67	3.94	16.42	4.41	3.16**
Change	16.92	4.44	17.20	4.87	.35
Endurance	12.19	4.19	12.63	5.19	.55
Heterosexuality	14.94	4.97	14.34	5.39	.68
Aggression	9.10	4.17	10.59	4.61	1.91*

*$p < .10$ **$p < .01$ ***$p < .001$

tended to be lower on Autonomy and Aggression (p between .05 and 10). All other comparisons yielded "t's" of less than 1.0. The findings for the group seem to confirm "clinical" impressions of the Student Nurse personality as less assertive than the usual female college student. The higher score of the nurses on Nurturance has an obvious face validity.

Table 2 gives the mean raw scores for each of the groups on the six PPS scales. Group means are given for the two combination scores, DSA and ADA, and for the D-R ratio. It can be seen here that all differences are in the expected directions. Contrasting the Rebellious group with the three Dependency groups, the Rebellious group is lower on Deference, Succorance and Abasement, and higher on Autonomy, Dominance and Aggression. The combination and ratio scores are also as expected. Table 2 also contains the "t" values for comparisons between the Rebellious group and each of the three Dependency groups. Among the individual scales, the Abasement and Autonomy scales seem to be most effective, particularly in distinguishing between the Rebellious and the Conformist and Dependent groups. The combination scores and the D-R ratio yield differences significant below the .05 level in every comparison between the Rebellious and the three Dependency groups. The more conservative two tailed test was used in all comparisons despite the fact that direction of the differences was implicit in the hypotheses. The fact that all significant differences are in the expected direction adds some weight to the results.

TABLE 2. GROUP MEANS AND SIGNIFICANCE OF THE DIFFERENCE ("t") BETWEEN THE REBELLIOUS GROUP AND OTHER GROUPS ON PPS MEASURES

	Rebellious	Submissive		Conformist		Dependent	
	M	M	t	M	t	M	t
Deference	10.1	14.2	2.24**	13.1	1.85*	13.5	1.85*
Succorance	10.6	13.3	1.33	14.6	1.86*	15.0	2.06*
Abasement	15.4	18.3	1.57	19.4	2.58**	20.8	3.69***
Autonomy	14.6	11.3	1.88*	9.4	3.40***	10.6	2.36**
Dominance	12.4	8.9	2.22**	11.3	.48	9.0	1.54
Aggression	10.0	9.2	.53	7.4	1.11	8.9	.55
DSA	140.4	163.3	5.61***	164.1	2.95**	170.1	5.27***
ADA	150.4	133.0	2.33**	130.1	3.15***	131.1	2.65**
D-R	.95	1.25	3.12***	1.28	3.00**	1.31	4.27***
N	7	10		7		8	

*p < .10 **p < .05 ***p < .01

Combinations of PPS scales were fairly effective in distinguishing between Dependent and Rebellious subjects. However, we cannot assume validity for other areas of personality. A group of seven subjects, who were characterized as "Hostile" by their peers, had the lowest mean of all the groups on the Aggression scale. A group of nine subjects characterized as "prejudiced" (toward minorities) had the highest group mean on this scale. It may be that the social undesirability of items on the Aggression scale is more obvious when paired with items on other scales. In all groups the mean raw score on Aggression was lower than on any of the other scales. In a study by Navran and Stauffacher[1], student nurses rated Aggression as the least desirable of the 15 PPS variables.

SUMMARY

A group of 63 student nurses was compared to Edwards' group of 749 female college students on the Edwards Personal Preference Schedule. Groups of Rebellious, Submissive, Conforming, and Dependent subjects were selected from the group of student nurses on the basis of peer ratings. The latter three groups were considered to fall into the general trait category of "Dependency". The Rebellious group was expected to be higher on the Autonomy, Dominance, and Aggression

[1]Leslie Navran and James C. Stauffacher, Personal Communication, March, 1958.

scales of the PPS and lower on the Deference, Succorance, and Abasement scales, than the three Dependency groups.

The student nurses were significantly higher on the Abasement and Nurturance scales and lower on the Autonomy, Dominance and Aggression scales of the PPS, than the general female college group. The Rebellious group of student nurses were significantly higher on the combination of Autonomy, Deference, and Abasement scales, than the three Dependent groups, and significantly lower than these three groups on the combination of the Deference, Succorance and Abasement scales. A Dependency-Rebelliousness (D-R) ratio was suggested for the measurement of this personality dimension from the PPS.

References

1. Bernardin, A. C. and Jessor, R. A construct validation of the Edwards Personal Preference Schedule with respect to Dependency. *J. consult. Psychol.*, 1957, *21*, 63-67.
2. Edwards, A. L. *Edwards Personal Preference Schedule, Manual.* New York: The Psychological Corp., 1954.
3. Murray, H. A. *Explorations in Personality.* New York: Oxford University Press, 1938.
4. Zuckerman, M. and Grosz, H. J. Suggestibility and Dependency. *J. consult. Psychol.*, in press.
5. Zuckerman, M., Norton, J. and Sprague, D. S. Acquiescence and Extreme Sets and their role in tests of Authoritarianism and Parental Attitudes. *Psychiat. Res. Reports*, in press.

MANIFEST ANXIETY AND MOTHER-SON IDENTIFICATION*

LAWRENCE H. STEWART

University of California, Berkeley

Introduction

The purpose of this study was to explore certain relationships between scores on the Taylor Manifest Anxiety Scale (MA) and mother-son identifications, defined operationally in terms of discrepancies between the Q-sorts of mothers and sons. Do boys who have small discrepancies between their self-ideals and their mother's ideals for them, for example, have low MA scores?

The nature of the psychological continua underlying the MA scale still remains to be determined but presumably it is a measure of the Hullian construct "drive." In conditioning experiments involving responses which have a high probability of occurring, such as in eyelid movements, performance of subjects with high MA scores is superior to that of subjects with low scores[6]. In more complex learning situations which involve competing responses, high scorers perform less adequately than do those with low scores[7]. These findings are consistent with prediction based on the construct of drive. Studies have demonstrated, however, that scores on the MA scale are highly correlated with scores on various measures of neuroticism[1, 2]. Jessor and Hammond[4] criticized the scale because it was devised as a measure of anxiety rather than as a measure of some aspect of drive. They pointed to some of the methodological difficulties in equating drive and neuroticism. Before the scale can be accepted as a measure of drive, alternative explanations for findings in learning experiments such as those indicated above must be ruled out. Obviously, the data presented in this brief report will not resolve these issues. Yet, accumulation of such data will eventually lead to a better understanding of what the scale is measuring.

*Data for this report were obtained as part of a larger study which was supported by a grant from the Research Committee, College of Education, University of California.

Method

The MA scale and a Q-sort were administered to a group of 92 adolescent boys. The MA scale was administered without the filler items which according to McCreary and Bendig[5] do not affect appreciably the reliability of the scale. The boys were asked to describe themselves on the Q-sorts in three ways: first, as they perceived themselves; second, as they would like to be; and third, as they thought their mothers would like them to be. Fifty-four mothers of these boys sorted the cards on two bases: as they perceived their sons and as they would like their sons to be.

The boys included in the sample were all juniors and seniors who were enrolled in the 1955 Summer Demonstration School of the University of California. All boys were Caucasian and were from sociologically intact homes, *i.e.*, both father and mother were living in the home. On enrolling in the school, students were arbitrarily assigned to rooms for testing. All boys in a room who met the sampling criteria were included. Rooms were selected for each period throughout the day. The mothers in the sample were those who could be reached by telephone and were willing to participate in the study. Mothers not included lived a considerable distance from the University, were on vacation for the summer, or were unable to arrange a convenient interview time because of work schedules.

The Q-sort contained 76 items obtained from sentence completions made by a group of students similar in age to those included in this study. Items were selected according to the categories used by Jersild[3] in classifying self-descriptions of children and adolescents. These categories were inner strength and drive, self-control, social attitude, relationships with opposite sex, intellectual abilities, home and family, physical characteristics, special talents, religion, and school relationships. For boys, the median test-retest reliability of the Q-sort over a two-week period was .71 for the self-sort and .68 for the ideal-sort.

Findings

Correlations between MA scores and discrepancies between the various combinations of sorts are shown in Table 1. Although all of the correlations were positive, only three were statistically significant at or above the .05 level. MA scores were significantly related to discrepancies between the boys' self-sorts and their ideal sorts (self-satisfaction), to discrepancies between their self-sorts and their perceptions of their mothers' ideals for them (meeting the perceived mother-ideal), and to

Table 1. Correlations Between Manifest Anxiety Scores and Discrepancies Between Sorts

Discrepancy	N[1]	Correlation
1. Boy's self and boy's ideal	81	.279*
2. Boy's self and his perception of his mother's ideal for him	73	.251*
3. Boy's ideal-self and his perception of his mother's ideal for him	72	.126
4. Mother's perception of her son and her ideal for him	53	.042
5. Boy's self and mother's perception of him	53	.105
6. Boy's ideal-self and mother's perception of him	52	.093
7. Boy's perception of the mother's ideal for him and her perception of him	52	.234
8. Boy's self and his mother's ideal for him	54	.489**
9. Boy's ideal-self and his mother's ideal for him	53	.218
10. Boy's perception of the mother's ideal for him and his mother's ideal for him	53	.185

[1]N's vary because complete data not obtained for all subjects.

*Significant at .05 level

**Significant at .01 level

those between boys' self-sorts and their mothers' ideals for them (meeting the mother-ideal). The last correlation was significant beyond the .01 level.

Analysis by inspection of the intercorrelations between the various measures of identification reported elsewhere (manuscript in press) indicated that they tended to fall into four major clusters. These clusters were not independent of each other, however; a variable in one might be highly correlated with one variable in another, yet independent of other variables in that cluster. These relationships were shown as secondary clusters. This analysis suggested the following types of mother-son identifications:

1. A realistic attempt to meet the mother's ideals, which leads to mother-satisfaction with her son.
2. A construal by the boy of the mother's reality perception of him as a mother-ideal, which also leads to mother-satisfaction.
3. An attempt by the boy to meet his perception of the mother's ideal for him, which leads to self-satisfaction. (This cluster was highly correlated with actually meeting the mother's ideal.)
4. An acceptance of the mother's ideal, which is closely related to his awareness of what the mother's ideal is.

All of the significant correlations between discrepancies and MA scores were in cluster three. Apparently the crucial relationships so far as MA is concerned were the boy's satisfaction with himself and the agreement between his self-perception and his mother's ideal for him—not the fact of accuracy in his perception of the mother's ideal or of his mother's satisfaction with him.

Summary

The purpose of this study was to investigate the relationship between Taylor MA scores and certain aspects of mother-son identifications as determined by discrepancies between Q-sorts made by a group of mothers and sons. All of the correlations between MA and discrepancies were positive although only three were statistically significant. Boys with low MA scores were satisfied with themselves, met their perception of their mother's ideal for them, and perceived themselves as being the kind of person their mother actually wanted them to be.

References

1. Brackbill, G. and Little, K. B. MMPI correlates of Taylor Scale of Manifest Anxiety. *J. consult. Psychol.*, 1954, *18*, 433-436.
2. Davids, A. Relations among several measures of anxiety under different conditions of motivation. *J. consult. Psychol.*, 1955, *19*, 275-279.
3. Jersild, A. T. *In search of self*. New York: Columbia University Press, 1952.
4. Jessor, R. and Hammond, K. R. Construct validity and the Taylor Anxiety Scale. *Psychol. Bull.*, 1957, *54*, 161-170.
5. McCreary, J. P. and Bendig, H. W. A comparison of two forms of the MA scale. *J. consult. Psychol.*, 1954, *18*, 206.
6. Taylor, Janet A. and Spence, K. W. Anxiety and strength of the USC as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, *42*, 183-188.
7. Taylor, Janet A. and Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, *44*, 61-64.

THE PSYCHASTHENIC AND HYPOMANIC SCALES OF THE MMPI AND UNCERTAINTY IN JUDGMENTS

ALBERT V. GRIFFITH,[1] HARRY S. UPSHAW,[2] AND RAYMOND D. FOWLER

University of Alabama

PROBLEM

In the validation of the clinical scales of the MMPI, the major criterion was the "... prediction of clinical cases against the neuropsychiatric staff diagnosis ..." (1, p. 6). The psychiatric constructs underlying these clinical scales assume a syndrome of behavioral characteristics; *e.g.*, psychasthenic individuals (identified by MMPI Pt scale) are said to be marked by vacillation, excessive doubt, worry, lack of confidence, and mild depression (1, p. 20; 6, p. 81). Hypomanic individuals (identified by MMPI Ma Scale) are said to be active, enthusiastic, confident, aggressive, and expansive (1, p. 21; 6, p. 167). Obviously, these two types of individuals should differ in the degree of confidence with which they make even simple judgments. This is the hypothesis tested in this study.

As a simple, objective measure of confidence in judgments, third-category or "doubtful" judgments in a psychophysical weight discrimination experiment were used. Much evidence has accumulated that the third-category measures confidence more than sensitivity. Woodworth (7, p. 65), for example, suggests that the frequency of third-category judgments is a measure of the "take a chance" attitude of the subject. If the subject is willing to take chances and follow the lead of slight impressions he will make a judgment while the more cautious subject will report doubt. Kellogg (2, p. 6) infers from this argument that the frequency of third-category judgments furnishes a better measure of a subject's confidence than it does of his discrimination.

With the measure of confidence defined, the experimental hypotheses may now be stated more exactly: (a) Individuals with a psychasthenic profile on the MMPI will give significantly more doubtful judgments than individuals with a hypomanic profile in a psychophysical weight discrimination experiment; and (b) Individuals with no T-score above 60 or below 30 on any of the MMPI scales (a "dull" profile) will give a number of doubtful judgments intermediate to the psychasthenic and hypomanic groups.

METHOD

Subjects. The subjects were male students entering the University of Alabama in the Fall semester, 1956, and the Spring semester, 1957. All had been given a battery of tests which included the MMPI. A total of approximately 1,200 MMPI profiles were screened. In selecting subjects for the experiment, none were chosen who had T-scores of 70 or above on any of the validity scales (L, F, or K). Those students who were not excluded by this condition, and who had "7' 2" profiles (the highest T-score, equal to or greater than 70, on the Pt scale, and the next highest T-score on the D scale) were classified as psychasthenic. Assignment to the hypomanic group was on the basis of "9' 4" profile (the highest T-score, equal to or greater than 70, on the Ma scale, and the next highest T-score on the Pd scale). Finally, a control group was selected on the basis that all of their T-scores fell between 30 and 60.

The total number of students selected by the above criteria was 28, 30, and 36 respectively for the psychasthenic, hypomanic, and control groups. Each individual was contacted by a letter from the office of the Dean of Men which requested his cooperation in the study. The result was that 16 subjects classified as psychasthenic, 14 as hypomanic, and 15 as control agreed to be tested.

Experimental task. In the psychophysical aspect of the study, the method used was that of constant stimulus differences. The experimental task was that of weight dis-

[1]Presently associated with Bryce Hospital, Tuscaloosa, Alabama.
[2]Presently with the Educational Research Corp., Cambridge, Mass.

crimination. A series of eight weights ranging from 92 to 108 grams with an interval of two grams between stimuli was used. The standard weight was 100 grams. The subject lifted the weights simultaneously, giving a judgment that either the weight in his right hand seemed heavier, the weight in his left hand seemed heavier, or he was doubtful as to which weight was heavier. By a randomizing procedure, the standard weight was placed equally often in the subject's right and left hand. Each subject gave a total of 144 judgments.

Results

The mean number of doubtful judgments for the psychasthenic, hypomanic, and control groups was 41.9, 20.1, and 31.0 respectively. Bartlett's test[6, pp. 206, 207] demonstrated that the variances of the judgments were significantly different ($\sigma^2_{PT} = 1013.8$; $\sigma^2_{MA} = 223.1$; $\sigma^2_{CONTROL} = 604.3$). It has been established that such violations of the parametric assumptions of the analysis of variance lead to too many rejections of the null hypothesis[3, p. 255]. To guard against this error, the Kruskal-Wallis procedure[4, pp. 184-193], a non-parametric analogue to the analysis of variance, was used to test for differences among the groups. The Kruskal-Wallis procedure employs for this purpose a statistic "H", distributed as Chi-square.

Table 1. Summary of the Kruskal-Wallis Procedure for the Data in the Present Study

Group	Number of Subjects	Sum of Ranks	"H"
Pt	16	459.0	
Ma	14	232.5	
Control	15	343.5	6.331*

*Sig. at the 5% level.

Table 1 summarizes the Kruskal-Wallis procedure for the data in the present study. As noted in Table 1, the "H" is significant at the 5% level of confidence. It is inferred, therefore, that a significant difference in the frequency of doubtful judgments exists among the groups, thus verifying the first hypothesis. An inspection of the means of the three groups verifies the second hypothesis. This behavioral correlate for the two MMPI scales provides, therefore, an indirect validation of the psychiatric constructs of psychasthenia and hypomania as measured by the MMPI.

Summary

This paper studied the relation of the psychasthenic and hypomanic scales of the MMPI to uncertainty in psychophysical judgment. The following predictions were made: (a) Individuals with a psychasthenic profile on the MMPI will give significantly more doubtful judgments than individuals with a hypomanic profile in a psychophysical weight discrimination experiment; and (b) individuals having a dull profile on the MMPI will give a number of doubtful judgments intermediate to the psychasthenic and hypomanic groups. Both predictions were verified.

References

1. Hathaway, S. R. and McKinley, J. C. *Minnesota multiphasic personality inventory manual.* New York: Psychol. Corp., 1951.
2. Kellogg, W. N. An experimental evaluation of equality judgments in psychophysics. *Arch. Psychol.*, 1930, No. 112.
3. McNemar, Q. *Psychological statistics.* New York: John Wiley, 1955.
4. Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.
5. Snedecor, G. W. *Statistical methods applied to experiments in agriculture and biology.* Ames, Iowa: Iowa State College Press, 1940.
6. Welsh, G. S. and Dahlstrom, W. G. (Ed.) *Basic readings on the MMPI in psychology and medicine.* Minneapolis. U. of Minn. Press, 1956.
7. Woodworth, R. S. Prof. Cattell's psychophysical contributions. In The psychological researches of James McKeen Cattell. *Arch. Psychol.*, 1914, *30*, 60-74.

COMPARABILITY OF MMPI CARD AND BOOKLET FORMS FOR PSYCHIATRIC PATIENTS[1]

ALBERT ROSEN
State College of Washington

WILLIAM M. HALES
U. S. Public Health Service

AND

ROLAND M. PEEK
Hastings (Minn.) State Hospital

Introduction

The original form of the MMPI, the card (individual) form, is now used much less commonly than the newer booklet (group) form. The card form is still popular, however, and is especially useful with patients who are aged, physically handicapped, poorly educated, or confused. In psychiatric settings the scores from both the card and booklet forms usually have been treated identically in clinical practice and research. It is important, therefore, to study systematically the comparability of scores from the two forms for psychiatric patients.

Procedure

The MMPI booklet form was administered to every other patient (with the exception of a few confused patients) admitted to the psychiatry section of theMinneapolis V. A. Hospital, until a sample of 75 was obtained. The scores from this group were compared with those from a sample of 250 patients from the same service who had taken the card form of the MMPI previously, and for whom complete MMPI statistics were already available from another study (2). This latter sample is representative of all new, white, male, testable veterans in the hospital over a two-year period.

The diagnostic characteristics of the two samples are shown in Table 1. It is apparent that the two samples are quite similar from the standpoint of diagnosis, the X^2 of 3.00 ($p > .30$) indicating that any differences obtained are not reliable.

Table 1. Diagnoses of Cases Administered the Card and Booklet Forms

Diagnosis	Percent in sample administered:	
	Card Form (N = 250)	Booklet Form (N = 75)
Neurosis	56	48
Psychosis	23	23
Personality disorder	12	20
Other diagnoses	9	9
Total Percent	100	100

Results

In Figure 1 are shown the mean scale scores and in Table 2 are listed the summary statistics for the two samples and the results of tests of significance of the difference between each pair of means. No significant differences between means were found. Moreover, *F*-tests indicated that no differences in variances were significant at the .05 level. These results are in essential agreement with those obtained by Cottle (1) for college students and by Wiener (3) for non-disabled veterans at a counseling center. Descriptive data on the diagnoses of individual cases who took the booklet form were no longer available when the final analyses were made for this report, so that it was impossible to compare the scores of subgroups on the two forms.

[1]This study was done at the Minneapolis VA Hospital.

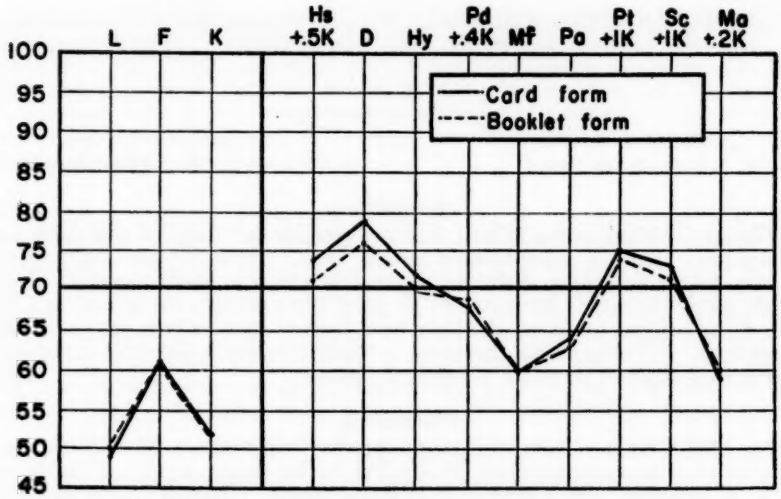


FIG. 1. MEAN MMPI SCORES FOR PSYCHIATRIC PATIENTS TAKING THE CARD AND THE BOOKLET FORMS.

TABLE 2. MMPI SCORES OF PATIENTS ADMINISTERED THE CARD AND BOOKLET FORMS

MMPI Scale	Card Form		Booklet Form		Critical Ratio
	M	SD	M	SD	
L	3.66	2.09	3.93	2.44	.87
F	7.64	5.41	7.39	5.54	.36
K	13.45	5.44	13.37	5.59	.11
Hs + .5*K*	20.71	6.55	19.41	7.00	1.43
D	28.71	7.35	27.72	6.79	1.09
Hy	28.42	6.56	27.69	7.75	.74
Pd + .4*K*	26.71	5.80	27.07	5.74	.45
Mf	25.49	5.67	25.63	5.41	.19
Pa	12.52	4.13	12.49	3.65	.05
Pt + 1*K*	35.25	7.98	34.27	7.31	1.00
Sc + 1*K*	34.24	9.66	32.85	8.95	1.15
Ma + .2*K*	20.66	4.99	21.09	5.06	.38

SUMMARY

A random sample of 75 patients in the psychiatry service of a VA General Medical and Surgical Hospital was administered the *booklet* form of the MMPI. Another essentially random sample of 250 patients from the same service was administered the MMPI *card* form. When the two samples were compared on each scale mean and variance, no significant differences were found. MMPI scores obtained from either the card or booklet form may in general be treated as alternate forms in the type of population sampled in this study. Further research is needed on the comparability of the two forms for population subgroups.

REFERENCES

1. COTTLE, W. C. Card versus booklet forms of the MMPI. *J. appl. Psychol.*, 1950, *34*, 255-259.
2. ROSEN, A. Development of some new Minnesota Multiphasic Personality Inventory scales for differentiation of psychiatric syndromes within an abnormal population. Ph.D. thesis, 1952, U. Minnesota. Microfilm of complete manuscript, University Microfilms, Ann Arbor, Mich., Publ. No. 4350.
3. WIENER, D. N. Differences between the individual and group forms of the Minnesota Multiphasic Personality Inventory. *J. consult. Psychol.*, 1947, *11*, 104-106.

A NOTE ON THE EFFECT ON SELF DESCRIPTION OF COMBINING ANXIETY AND HOSTILITY ITEMS ON A SINGLE SCALE

SAUL M. SIEGEL — *Topeka State Hospital*

AND

MARVIN J. FELDMAN — *University of Buffalo*

In a recent article[1] the present authors tested the hypothesis that the total item composition of a test affects responses to the component parts of the test, *i.e.*, that related variables like anxiety and hostility would be especially prone to show such inter-acting effects. Three related personality inventories were used: (a) The Taylor Manifest Anxiety Scale (MAS)[3] - 50 items pertaining to anxiety plus 175 buffer items; (b) The Siegel Manifest Hostility Scale (MHS)[2] - 50 items pertaining to hostility plus the same 175 buffer items; and (c) Combined form—50 MAS, 50 MHS, plus 125 of the same buffer items. The results indicated that while the MAS scores were not affected by inclusion of hostility items, the MHS scores were reduced by the inclusion of anxiety items. The present study was designed to cross validate these findings with a more refined approach and a larger number of subjects.

Procedure

As in the previous study, the limiting case of self ratings was used. In this instance, the number of related personality inventories was increased to five with 100 items in each scale: Form I — The Taylor Manifest Anxiety Scale (MAS) of 50 items followed by The Siegel Manifest Hostility Scale (MHS) of 50 items; Form II — the MHS followed by the MAS; Form III — Each MAS item alternating with an MHS item; Form IV — The MAS interspersed with neutral items; Form V — The MHS interspersed with 50 neutral items. The neutral items used in forms IV and V were based upon judges' ratings of the buffer items from the Taylor MAS. A few items were added to the buffer pool to arrive at 50 items that four of the six judges agreed were neutral in content.

The 374 subjects were mainly freshmen and sophomores enrolled in two large introductory classes. The tests were administered as part of the regular class work. The proportion of males to females averaged 42 to 35 respectively for those taking each form of the test with little variability between groups. The average age of the subjects was 18.7 years, again with little variability between the groups taking the different forms. One-fifth of a class each took a different form of the test. The members of the class, however, were not aware that they were taking different forms.

Results and Conclusions

Table 1 presents the means and standard deviations that were produced by the two scales when combined in the various forms used. They are quite similar to those reported previously[1].

Table 1. Means and Standard Deviations on the MAS and the MHS

Form	MAS N	MAS M	MAS σ	MHS N	MHS M	MHS σ
I	80	13.0	8.0	80	13.1	7.0
II	74	14.2	8.5	74	12.9	5.5
III	74	13.0	7.6	74	12.9	6.6
IV	72	13.9	8.9			
V				74	13.7	5.2
Total	300			302		

Table 2 presents the results from the analysis of the MAS and MHS scores. Neither of the F ratios was significant. Thus it appears that MAS scores are not

TABLE 2. ANALYSIS OF VARIANCE OF MAS AND MHS SCORES

Source	MAS Scores df	E sqs	ms	F	MHS Scores df	E sqs	ms	F
Among S	3	89.15	29.71	.43	3	31.15	10.38	.27
Within S	296	20,399.69	68.91		298	11,252.82	37.76	
Total	299	20,488.84			301	11,283.97		

affected by the inclusion of hostility items and that the MHS scores are not affected by the inclusion of anxiety items. Also neither scale score seems to be affected any differently whether it is neutral items that are included or items from the other scale. These results suggest that combining subscales or adding neutral items has no effect on the responses to either scale.

REFERENCES

1. FELDMAN, M. J. and SIEGEL, S. M. The effect on self description of combining anxiety and hostility items on a single scale. *J. clin. Psychol.*, 1958, *14*, 74-77.
2. SIEGEL, S. M. The relationship of hostility to authoritarianism. *J. abnorm. soc. Psychol.*, 1956, *52*, 368-372.
3. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, *48*, 285-290.

SCORES OF ALCOHOLICS ON THE SIXTEEN PERSONALITY FACTOR QUESTIONNAIRE[1]

NICHOLAS DE PALMA AND HUGH D. CLAYTON

Davidson County Hospital, Nashville, Tenn.

PROBLEM AND PROCEDURE

The present study attempts to discriminate between the "alcoholic personality" and the general population using the Sixteen Personality Factor Questionnaire (16 P. F.).(1)

The data for this investigation are based on profiles drawn at random from the hospital records of alcoholic in-patients, all residents of Davidson County, Tennessee, committed to the hospital for a period of sixty days by the county court. Commitment results after all other resources have failed and represents a relatively complete failure on the part of the individual to adapt to his problems.

The sample consists of sixty-nine male subjects between the ages of thirty and fifty-five, whose average education is approximately nine years. Reading comprehension was required for the examination.

The hypothesis is that our obtained means do not differ significantly from those of the general population reported by the authors of the test for each of the sixteen factors. The tests were two tailed since each direction of deviation has a separate and distinct meaning, with Alpha set at .05. The authors of the test have published norms for both males and females separately. Naturally, our results are tested against their male norms.

[1]This study was made possible by a grant from the Tennessee Alcoholic Commission. Acknowledgment is due to its chairman, Judge Henry F. Todd, and to the superintendant of the Davidson County Hospital, Dr. Joseph J. Baker.

Results

The statistical results are summarized in Table 1. The first of the two personality attributes listed for a given scale is represented by a low score, i.e., one below the

Table 1. Statistical Analysis of Raw Scores of Alcoholics (N = 69) vs. General Population (N = 549)

Factors		Alcoholic		General Pop.			
		M	SD	M	SD	"t"	p
A	Aloof *vs.* Sociable	8.49	2.8	7.5	3.1	2.5	.02
B	Dull *vs.* Intelligent	5.88	2.0	8.8	1.7	13.27	.001
C	Emotional *vs.* Mature	14.30	3.5	19.2	2.7	10.7	.001
E	Submissive *vs.* Aggressive	12.57	3.5	15.2	3.3	4.0	.001
F	Silent *vs.* Talkative	12.10	2.9	16.2	3.2	10.25	.001
G	Undependable *vs.* Conscientious	12.41	2.8	13.8	3.1	3.56	.001
H	Timid *vs.* Adventurous	12.51	3.7	14.0	4.7	3.04	.01
I	Tough *vs.* Sensitive	8.59	2.5	6.4	3.1	6.6	.001
L	Trustful *vs.* Suspecting	8.30	3.1	6.9	3.2	3.41	.001
M	Conventional *vs.* Bohemian	11.75	3.0	9.2	3.2	6.2	.001
N	Awkward *vs.* Sophisticated	11.84	2.4	12.1	2.7	.76	NS
O	Confident *vs.* Anxious	10.72	4.0	7.3	3.2	6.84	.001
Q^1	Conservative *vs.* Radical	9.23	2.3	10.1	2.9	2.90	.01
Q^2	Dependent *vs.* Self-Sufficient	10.01	2.9	9.5	3.5	1.34	NS
Q^3	Unsure *vs.* Controlled	11.97	3.1	11.1	2.8	2.42	.02
Q^4	Composed *vs.* Tense	11.39	5.0	9.0	4.4	4.19	.001

standardization mean, the second by a high one. For further description of the categories, refer to the test manual. Inspection of our results reveals the fact that the alcoholics differ significantly from the general population on fourteen of the sixteen scales, of which several attain very high levels of significance. As one might suspect, our alcoholics measure markedly unstable and immature emotionally. They obtain extremely low scores on the "emotional vs. mature" scale, significant beyond the .001 level and constituting the greatest deviation of all. In addition, our sample of alcoholics have not made full use of their inherent intellectual capacities. They are further described as being silent, submissive, sensitive, suspecting, and measuring quite low in superego strength. They are socially undependable, Bohemian, anxious and insecure, and high in ergic tension.

Discussion

Forty-five of these same alcoholics were administered the Wechsler Adult Intelligence Scale and their mean IQ's are reported as follows: Verbal 95, Performance 90, Full Scale 92. From these scores one could hypothesize that alcoholics fall within the average range on factor "B" of the "16 P. F. Test". This hypothesis is not born out; the mean score for alcoholics falls significantly below the average range. Concerning these findings, Cattell states in a personal communication: "You will notice that the intelligence test, being incorporated in a questionnaire is given under rather different conditions than usual, in that it corresponds to an untimed test, and also puts greater emphasis on initiative and on adjusting to the unusual situation of encountering such items among ordinary questionnaire items. Accordingly, our experience has been all along that this is more a measure of effective use of resources of intelligence than of intelligence itself. Of course there is also the fact that this B measurement is made on decidedly fewer items than usually enter into an intelligence test, so that it cannot be as reliable as a test specifically shaped for that purpose. However, I think on a group of the size which you have, the lower score here than on a formal Wechsler is best interpreted as an ineffectiveness in applying intelligence in a less supervised situation."[2]

[2]Cattell, R. B. Personal communication. February 4, 1958.

With the exception of his comment regarding the time factor, the authors concur with Cattell's statement. Certainly, much of the data gleaned from the social histories point to the fact that alcoholics have squandered many of their skills and potentials. Factor "B", as an untimed test of intelligence, is more comparable to the Verbal scale of the Wechsler than to the Performance. Yet, the Verbal scale is essentially untimed also. The discrepancy between Factor "B" and the Wechsler cannot be explained as due to the time variable.

Emotional immaturity seems to be the basic core of the alcoholic personality. The alcoholic's behavior is governed by the "pleasure principle" rather than by the "reality principle". He has difficulty forestalling, postponing, or giving up immediate gratification of desires in order that he might realize more distant goals. The capacity for inhibiting basic needs and impulsive emotional release is poorly developed. Low tolerance for frustration, stress, tension, and duress are characteristic features of the alcoholic makeup. Though the causes for a person's drinking are not clearly verbalized, still the patient invariably relates his drinking to "nervousness" and tension. Hence, there seems to be a common vein running through the test results. This is best described as a "stress syndrome" consisting of emotional immaturity, high ergic tension, sensitivity, suspicion, anxiety, and insecurity. One is not certain as to why drinking first started; but the contemporary motive for continued indulgence is the need to allay tension and tremor. Such seems to be the precipitating "cause" of chronic alcoholism. It seems almost banal to conclude that stress begets alcoholism and alcoholism enhances stress. Yet, that seems to be the dynamic cycle.

Another personality pattern to be evidenced is that relating to social deviation. Strong elements of Bohemianism along with low superego strength attest to the alcoholic's egocentricity. This seems to be a symptom over and above alcoholism per se. Apparently alcoholism still remains a form of sociopathology.

Finally, the alcoholic tends toward warm and sociable interpersonal relations. His social role is more likely to be characterized by submissiveness and desurgency rather than aggression and enthusiasm.

Summary

Sixty-nine alcoholic in-patients were administered the "16 P. F. Test" and their scores were tested statistically against Cattell's norms using two-tailed "*t*" tests. At or beyond this level, fourteen of the sixteen factors showed significant differences. The alcoholic personality is characterized by squandered intellectual potentials, low tolerance for stress, sociopathology, and submissive sociability.

Reference

1. Cattell, R. B., Saunders, D. R. and Stice, G. *Handbook for the Sixteen Personality Factor Questionnaire.* Institute for Personality and Ability Testing, Champaign, Illinois, 1957.

THE LEARY INTERPERSONAL CHECK LIST: A RELIABILITY STUDY

RENATE GERBOTH ARMSTRONG

East Moline (Ill.) State Hospital

The Leary Interpersonal Check List[5, 6] was developed to provide (a) a multilevel diagnosis which evaluates the individual personality, and (b) a diagnosis of the interpersonal structure of a group. Leary and his co-workers have investigated the stability of the Check List by means of test-retest correlations; these ranged from .64 to .83, averaging .78 for octant reliability, *i.e.* stability of octant scores[4]. However, this writer feels that the reliability should be further investigated by means of the internal consistency of the test. This might tell us if "the test is relatively free of chance errors of measurement and scores earned on it are stable and trustworthy"[2, p. 380]. It is the purpose of this study to determine the internal consistency of the Check List.

Probably the most appropriate statistic to tap this type of reliability is the Kuder-Richardson estimate of reliability. This formula avoids the pitfalls of the split half reliability correlations by the use of item statistics. It does require a power, rather than a speed, test which has items of equal difficulty. The Check List fulfills both of these requirements. However, a further requirement, that of homogeneity, *i.e.*, that the test taps one main trait, is somewhat dubious, for the Check List is set up to investigate eight aspects or octants of interpersonal behavior. Still, Leary[6] has found that intercorrelations between the items and the octants do exist, for in the before mentioned reliability study[4] octant intercorrelations ranged from .46 to .12. Therefore, it is felt that the Kuder-Richardson formula can still be used with this test.

The formula employed was $\left(\frac{n}{n-1}\right) \times \frac{(\sigma_t^2 - \Sigma pq)}{\sigma_t^2}$, which, according to Guilford[3], is considered the most accurate since other formulae tend to underestimate the reliability of the test investigated. Further, some of the alternate formulae have been considered only approximations of the above formula[2].

The samples reported here were collected for a previous study[1], and consisted of 50 "normal" males and 50 alcoholic males matched for age, education, and occupation. They were of average intelligence, mean IQ's 106.58 and 108.80, respectively. Also, none of the alcoholics were psychotic, but all were diagnosed Sociopathic Personality Disturbance, Addiction, Alcoholism. The 100 subjects were given the Leary Interpersonal Check List, on which each subject was asked to rate himself, his mother, his father, his wife, his ideal wife, and his ideal self. Therefore, six ratings were obtained from each of the two groups. The above given Kuder-Richardson formula was applied to these 12 ratings. Significance of each r_{tt} was tested for, as well as significance of a difference between each of the twelve r_{tt}'s.

The statistical results are shown in Table 1. It can be noted that each of the r_{tt}'s is significant beyond the .01 level. The range (.953 to .976) of but .023 points indicates that all 12 ratings have a most similar degree of internal consistency; also, testing of differences between all the r_{tt}'s revealed that none of them were significantly different. Therefore, both the "normal" and the alcoholic groups, as well as the six ratings yielded most similar Kuder-Richardson estimates of reliability.

The overall findings indicate that regardless of population (in this case "normal" or alcoholic) and regardless of what rating asked for (self, mother, father, etc.), the Leary Interpersonal Check List yields a highly significant internal consistency. Therefore, the test can be considered to be quite reliable, at least as far as the present investigation has gone.

TABLE 1. KUDER-RICHARDSON ESTIMATES OF RELIABILITY AND THEIR SIGNIFICANCE APPLIED TO SIX CHECK LIST RATINGS OF TWO GROUPS

Check List Ratings	r_{tt}	Significance
Normals (N = 50)		
Self rating	.953	.01
Mother rating	.967	.01
Father rating	.969	.01
Wife rating	.966	.01
Ideal wife rating	.958	.01
Ideal self rating	.953	.01
Alcoholics (N = 50)		
Self rating	.958	.01
Mother rating	.967	.01
Father rating	.976	.01
Wife rating	.972	.01
Ideal wife rating	.960	.01
Ideal self rating	.966	.01

SUMMARY

This study was undertaken in order to determine the reliability, *i.e.*, internal consistency, of the Leary Interpersonal Check List. Kuder-Richardson estimates of reliability were applied to six ratings obtained from a group of 50 "normal" and a group of 50 alcoholic males. These r_{tt}'s ranged from .953 to .976, all of them being highly significant, and no significant differences occurring between any of the twelve reliability coefficients. Judging from these results, therefore, the Leary Interpersonal Check List can be considered to have high internal consistency.

REFERENCES

1. ARMSTRONG, RENATE G. Personality structure in alcoholism. Unpublished doctoral dissertation, Univ. of Colorado, 1957.
2. GARRETT, H. E. *Statistics in psychology and education.* New York: Longmans, Green and Co., 1947.
3. GUILFORD, J. P. *Fundamental Statistics in psychology and education.* New York: McGraw-Hill Book Co., 1950.
4. LAFORGE, R. and SUCZEK, R. F. The interpersonal dimension of personality: III. The interpersonal check list. *J. Pers.*, 1955, *24*, 94-112.
5. LEARY, T. F. The theory and measurement methodology of interpersonal communication. *Psychiat.*, 1955, *18*, 147-161.
6. LEARY, T. F. *Multilevel measurement of interpersonal behavior.* Berkeley, Calif.: Psychological Consultation Service, 1956.

SELECTIVE AND DIFFERENTIALLY ACCELERATED INTELLECTUAL DYSFUNCTION IN SPECIFIC BRAIN DAMAGE

GRANVILLE C. FISHER

University of Miami

PROBLEM

Psychometric studies of brain damage have produced ambiguous results which may stem from indiscriminate grouping, without taking into consideration the specific type of pathology. Piotrowski's "organic personality" and Wechsler's "organic signs" are questionable. Perhaps there are intellectual and emotional commonalities among all subjects with organic mental disease, but some investigations [3, 4, 5] indicate that there also may exist significant differences attributable to the various types of brain damage.

This study hypothesizes that mental deterioration is unevenly accelerated in the various intellectual functions; and also that a specific brain disorder is selective as regards those functions which will most quickly succumb to, and those which will withstand the longest, pathological damage to cerebral tissue. *General paresis* has been selected as a clearly defined sub-group among organics to serve as the primary group. Comparisons will be made between this group and one diagnosed as *Korsakoff's psychosis* to test that part of the hypothesis referring to the selectivity of the disorder.

PROCEDURE

The Wechsler-Bellevue Scale was selected as the psychometric instrument. The files of all male patients in a state hospital with the diagnosis of general paresis were examined. All over 60 years of age were eliminated. Below 60 the age factor is controlled on the W-B if one makes use of the tables of means and standard deviations of subtest scores distributed by ages. Also eliminated were all who were mental defectives prior to infection; all chronic epileptics; all with language handicaps; and all who did not have unanimous staff diagnosis.

In the selected group all showed a Lange gold sol curve with three precipitations or better in the first zone, and the whole curve characteristically paretic; *i.e.*, the first zone highest, with the count diminishing in the middle and last zones. Positive Wasserman or Kahn was demanded. Using the criteria of Roe and Shakow [6] of "representative" and "non-representative" test results, all cases which were inaccessible or uncooperative were eliminated. Twenty-six paretic cases met these criteria. Sixteen cases with Korsakoff psychosis also met a similar set of criteria.

RESULTS

Table 1 presents data on the paretic and Korsakoff groups.[1] For our purposes, it appears more legitimate to use standard deviation scores based on Wechsler's tables, rather than the weighted scores, inasmuch as the effect of age would thus be taken into account. Since no such scores were provided for Vocabulary, the weighted mean of 10 and the S.D. of 3 were used.

TABLE 1. MEAN STANDARD SCORES OF PATIENTS WITH GENERAL PARESIS AND KORSAKOFF'S PSYCHOSIS

		Inf.	Com.	Ari.	Dig.	Sim.	Voc.	P.A.	P.C.	B.D.	O.A.	D.S.
GP	M	−.42	−1.17	−.67	−1.23	−1.49	−.62	−1.07	−.59	−.92	−.91	−1.00
	Rank	1	9	4	10	11	3	8	2	6	5	7
K	M	−.50	−.35	−.66	−.54	−1.23	−.44	−.57	−.21	−.66	−.16	−.92
	Rank	5	3	8.5	6	11	4	7	2	8.5	1	10

[1] More complete information on the individual subjects and the raw Wechsler-Bellevue data may be obtained on request from the author.

Table 2 indicates the mean deviation score of every subtest as below average. This may indicate that paretics are below average intelligence. But this does not confirm whether they were so prior to their disorder, or are so as a result of their disease. The raw data suggest that all were of at least average intelligence premorbidly.

TABLE 2. *t* SCORES OF SUBJECTS WITH GENERAL PARESIS*

S.D. Rank	Subtest	Voc.	P.C.	Ari.	O.A.	B.D.	D.S.	Com.	P.A.	Dig.	Sim.
1	Information	1.46	.91	.98	2.11	2.72	2.83	3.83	4.13	3.93	5.45
3	Vocabulary		.14	.22	1.26	1.32	2.12	2.54	3.05	3.06	4.08
2	Picture Com.			.32	1.69	1.69	2.16	2.96	2.70	2.21	3.88
4	Arithmetic				.89	1.18	2.00	1.96	1.92	2.59	3.34
5	Object Assem.					.06	.46	1.14	.81	.98	2.58
6	Block Design						.42	1.13	.83	1.07	2.73
7	Digit Symbol							.75	.46	1.07	2.33
9	Comprehension								.44	.23	2.13
8	Picture Arr.									.67	2.20
10	Digit Span										1.08
11	Similarities										

*Significance at the one per cent level of confidence = 2.787; five per cent level of confidence = 2.06.

A wide range of variation, from a low score of −3.26 to a high of 2.23, was found in the deviation scores. An analysis of variance shows the larger variance between the sub-tests rather than within them. The F of 2.53 is significant at the one per cent level of confidence. This suggests that some significant law of variation operates among the sub-tests.

Table 2 is a matrix of t scores of critical ratios of the differences between each sub-test and every other sub-test. These are arranged in order to create a minor matrix uncontaminated by any value below the one per cent level of confidence. This consists of six cells in the upper right hand corner, three horizontal and two vertical. Since these are the most discriminating, and include the widest possible range of functions, they allow us to devise an index characteristic of the paretic's comparative scoring on this scale.

The S.D. scores were translated into Z scores to eliminate the plus and minus signs. The two top-ranking sub-tests, Information and Vocabulary, were added and the sum multiplied by 1.5. From this was subtracted the sum of the three lowest ranking sub-tests, Similarities, Digit Span, and Picture Arrangement. The remainder was then divided by the product of the first half of the computation, and this then multiplied by 100 in order to remove the decimal point. The equation reads:

$$\frac{1.5\,(\text{Inf.} + \text{Voc.}) - (\text{Sm.} + \text{DS} + \text{PA})}{1.5\,(\text{Inf.} + \text{Voc.})} \times 100$$

This index is expressive theoretically of the per cent of deterioration of the second or low group of sub-tests beyond that of the first or high group. For the subjects as a whole the mean index is 16.6, with a S.D. of 11.5, a S.E. of 2.3, and a t score of 7.2 (1% = 2.787). The index for every individual in this group is positive with the exception of one. Only four cases are below one sigma.

The data in Table 1 reveal that the Korsakoff group has not suffered as great a loss of global intelligence as has the paretic group (Korsakoff total scale S.D. mean, −.57; Paretic, −.96). In both groups Similarities ranks lowest, and Picture Completion second from the top. There is no other direct correspondence. Rho between the two rank orders is .55. The equivalent r is approximately .568.

The analysis of variance for this group showed no significant F at the one per cent level. The assumption is that the differences between the sub-tests are not significant.

Discussion

Wechsler states [7, pp. 153 ff.] that "organic brain cases, with few exceptions, do consistently better on verbal than on performance tests." In this study, of the five lowest subtests in the paretic group, three are verbal. The mean Verbal IQ is 84, and Performance 85. Wechsler states again that "the greatest and most consistent falling off is on the Digit Symbol Test." In the rank order of the paretics, Digit Symbol is in the middle group, six tests being above and four below. Wechsler then states that even more diagnostic "is the organic's inability to do the Block Design test." In this study Block Design stands exactly in the middle of the rank order.

"The four most conspicuous signs of organic brain disease," says Wechsler [7, p. 161], are a "large discrepancy between Verbal and Performance in favor of the former, very low Blocks combined with even lower Object Assembly and very low Digit Symbol." Not one of these "most conspicuous signs" is found in the paretic group of this study. Hence, there seems to be some basis for concern regarding the heterogeneous character of over-generalized categories. The marked discrepancies between Wechsler's findings and those of this study indicate the desirability of avoiding grab-bag categories, and lend support to the earlier contention that there may be merit in distinguishing between types of organic brain damage in psychometric research.

Whereas the results of our study reverse Wechsler's "organic" signs on so many counts, using the same instruments they corroborate the findings of Roe and Shakow [6] on nearly every count, even though the instrument employed is different. Wechsler apparently used an indiscriminate, overly-generalized classification; Roe and Shakow [6] used a more clearly defined disorder.

Several psychometric investigations have reported that deterioration in the presence of paresis is general rather than differentiated [1, 2, 7]. Our data lend rather vigorous support to the contrary view; *i.e.*, that the effect of general paresis is a *selective* impairment. The *t* scores in Table 2 do not indicate a single sub-test which at the five per cent level is not significantly distant from at least one other sub-test on the basis of expected scores from normal individuals; that is to say, on the basis of standard deviation score differences. Object Assembly and Block Design are the only two not significantly distant from at least one other at the one per cent level.

The differences between the standard deviation scores for Information (rank 1) and Similarities (rank 11) were computed for each case. The minimum difference significant at the one per cent level is .55. Nineteen scored a difference higher than the minimum, four showed a positive trend, and three reversed the order. As a group phenomenon the selectivity aspect of the hypothesis is upheld, whereas on an individual basis these data do not support it completely.

The question which now arises is to what extent differential acceleration of *any* type of functional loss applies to the individual case. To answer this question the minimum mean deviation which is significant at the one per cent level was computed for every dimension in the sub-test matrix with a *t* score in excess of the critical minimum. There is not a single subject who does not show a significant difference in his scores between at least one pair of sub-tests. The range is from 1 to 11. Two individuals show only one significant difference score. Two show eleven. The mean is 6.31. If positive trends are added, the range is from three to eleven, with the mean at 8.53. The range of reversals is from zero to seven, with the mean at 2.46.

Summary and Conclusions

The hypothesis was tested that mental deterioration is unevenly accelerated in the various intellectual abilities, and that a specific brain disorder is selective as regards those functions which will most quickly succumb to, and those which will withstand the longest, pathological encroachments upon cerebral tissue. Twenty-six paretics and 16 Korsakoff cases were examined with the Wechsler-Bellevue Scale.

An index was devised which is expressive theoretically of the per cent of deterioration of the low group of sub-test scores beyond that of the high group. The

index was positive for every subject with the exception of one. Only four cases were below one sigma. This, together with eleven *t*-scores above the critical minimum is supportive of that part of the hypothesis which holds that deterioration is differentially accelerated in the brain-damaged individual.

Although paresis is not invariably selective, the preponderance of subjects in this study did manifest such selectivity. Every subject showed a significant difference in his scores between at least one pair of sub-tests.

Within the conditions and limitations of this research, and pending cross validation, the following conclusions are suggested:

1. The completely negative set of sub-test mean scores supports the view that patients with general paresis undergo significant alterations in intellective functional capacity; that in general a deleterious effect upon intellectual efficiency, and probably capacity as well, obtains; and that as a group characteristic there appears to be a selective impairment in specific abilities.

2. General paresis appears to affect certain intellective functions more adversely than others, which is reflected in a disparate rate of deterioration in most individual cases.

3. The two functions most adversely affected are verbal concept formation as reflected in Similarities, and rote memory as reflected in Digit Span. The function which is seemingly most resistive is remote memory, as reflected in Information and Vocabulary.

4. Vocabulary does not invariably "hold up" significantly better than other functions, and hence may not be universally utilized as a base for establishing premorbid intelligence. A test of information may be as adequate, or even more so.

5. "Organic brain damage" as a category is insufficiently discrete for psychometric diagnosis, and can lead to fallacious and erroneous evaluations.

6. The multidimensional theory of intelligence is supported in view of the consistent and significant disparities in ability displayed in these data.

7. General paresis affects "verbal" abilities to as great an extent as "performance" abilities, contrary to many past pronouncements. Hence, the dichotomizing of tests into "verbal" and "performance" with the implication of functional unity seems not to be justifiable, at least for this group.

Bibliography

1. Binet, A. and Simon, T. H. *The Intelligence of the Feeble-Minded* (Trans. by E. S. Kate). Baltimore: Williams and Wilkins Co., 1916.
2. Gilliland, A. R., Wittman, P., and Goldman, M. Patterns and Scatter of Mental Abilities in Various Psychoses. *J. gen. Psychol.*, 1953, *29*, 251-260.
3. Harrower-Erickson, M. R. Personality Changes Accompanying Cerebral Lesions. Rorschach Studies of Patients with Cerebral Tumor. *Arch. Neurol. Psychiat.*, Chicago., 1940, *43*, 859-890.
4. Hebb, D. O., and Penfield, W. Human Behavior After Extensive Bilateral Removal from the Frontal Lobes. *Arch. Neurol. Psychiat.*, Chicago, 1940, *44*, 521-538.
5. Hunt, H. F. A Note on the Clinical Use of the Hunt-Minnesota Test for Organic Brain Damage. *J. appl. Psychol.*, 1944, *28*, 175-178.
6. Roe, A., and Shakow, D. Intelligence in Mental Disorder. *Ann. N. Y. Acad. Sc.*, 1942, *42*, 361-490.
7. Wechsler, D. *The Measurement of Adult Intelligence.* 3rd Ed. Baltimore. Williams and Wilkins Co., 1944.

INTELLIGENCE TEST CORRELATES OF SHIPLEY-HARTFORD PERFORMANCE[1]

LLOYD K. SINES[2]

Fergus Falls State Hospital, Minnesota

PROBLEM

The present report deals with the validity of the Shipley-Hartford Retreat Scale as a brief paper-and-pencil measure of current intellectual functioning. Although the Shipley-Hartford scale was devised primarily as a scale to reflect intellectual deficit [20], and its subsequent evaluation has indicated limited utility as a measure of deterioration [2], several investigators have been interested in its usefulness as a measure of functioning intelligence. [8, 16] Table 1 summarizes some relationships between various brief measures of intelligence and the Wechsler Bellevue Scales.

TABLE 1. STUDIES REPORTING RELATIONSHIPS BETWEEN VARIOUS BRIEF, PAPER-AND-PENCIL, OR OTHER SIMILAR ESTIMATES OF INTELLECTUAL FUNCTIONING AND THE WECHSLER-BELLEVUE SCALES

Test	Investigator	Sample		Validity Coefficient
ACE	Merrill [16]	251	college students	.61
AGCT	Tamminen [22]	100	VA counseling clients	.83
Ammons FRPV	Allen, et al. [3]	51	VAH psychiatric patients	.86
Army Alpha	Rabin [17]	92	student nurses	.74
Goodenough	Jones [11]	40	aged males	.65
Immediate Test	Corsini [4]	100	prison inmates	.89
Kent EGY	Delp [5]	513	students (plus others)	.65
Kent EGY	Robinowitz [19]	50	hosp. psychiatric patients	.69
Kent (7 tests)	Knott, et al. [12]	100	hosp. patients (mostly psych.)	.87
Otis SA	Goldfarb [9]	108	adult males	.73
Progr. Matrices	Stacey [21]	172	subnormal adults	.68
Rorschach (4 factors)	Abrams [1]	400	VA MHC patients	.53
Shipley-Hartford	Garfield [8]	100	hospitalized veterans	.68
Shipley-Hartford	Wright [24]	134	psychiatric patients	.77

Wright [24] has stated that "The Shipley-Hartford test can be used as a rough estimate of functional intelligence," while Manson and Grayson [14] concluded that although "The Shipley test was not a valid test of deterioration . . . (it) was a valid test of intelligence." The latter investigators provided mental age and AGCT score equivalents of Shipley total raw scores, but Garfield [7] has found that Shipley MAs are too low, when compared with Wechsler scores, ". . . particularly for individuals in the lower half of the intelligence distribution." In an attempt to further evaluate the Shipley-Hartford scale as a brief measure of intelligence, data bearing on the problem were obtained from three sources and are reported below.

PROCEDURE AND SUBJECTS

The first sample (N = 99) was obtained by exhausting the psychological test files at a Veterans Administration Mental Hygiene Clinic (VA MHC) of all records of patients given both Wechsler-Bellevue and Shipley-Hartford tests. Records were selected only if the subject was a male veteran and had been given at least four of the verbal, and three of the performance subtests on the Wechsler. In fifty-three of the cases the Wechsler-Bellevue and the Shipley-Hartford had not been administered at the same time, and in these instances, the time interval between the two tests

[1]The author is indebted to Dr. H. G. Gilberstadt for assistance in collecting and recording the data from the VA Hospital, Ft. Snelling, Minnesota.

[2]Formerly of the Veterans Administration Mental Hygiene Clinic, Ft. Snelling, Minnesota.

ranged from two to eighty months. It should be noted that the sample obtained is neither a random nor necessarily a representative sample of the total clinic population since it was not standard practice to administer both Wechsler-Bellevue and Shipley-Hartford tests to all patients. For example, a patient may have been given a Wechsler if there was a question of organicity, or if he had done poorly on the Shipley-Hartford.

The second sample was obtained from the psychological test files on the psychiatric service of a VA GM&S hospital. The procedure here was similar to that employed at the VA MHC, and a total of 251 records was thus obtained. In each of the two samples, Shipley vocabulary, abstractions, and total raw scores were correlated with Wechsler verbal, performance, and full scale total weighted scores as well as with the respective IQs. A regression equation for predicting Wechsler full scale scores was derived for each sample using the Shipley variable (total raw score) showing the highest relationship to the Wechsler. A table for converting Shipley total raw scores to Wechsler full scale *weighted scores* was constructed for each sample.

The third sample of subjects was obtained in connection with a major research project involving psychiatric aides in a state hospital. Forty female and twenty male aides were randomly selected from a larger group meeting the following criteria: (a) having a civil service classification of Psychiatric Aide I, (b) having completed a one-year in-service training program, and (c) being available for research duties for a one-year period in the hospital. The aides were given a battery of psychological tests which included both the Shipley-Hartford and the civilian form of the AGCT. The three Shipley raw scores were correlated with both AGCT score and AGCT raw score. The regression equation between Shipley total raw score and AGCT score was derived and a conversion table was constructed.

Results and Discussion

Table 2 summarizes the essential data from the two VA samples, while Table 3 presents the comparable data from the state hospital psychiatric aide sample. Table 4 contains the correlation coefficients between Shipley-Hartford scores and the other two intelligence tests used in the samples studied. It may be seen from Table 4 that the relationship between the Shipley-Hartford and the other two measures of intellectual functioning compares very favorably with those reported between various other measures and the Wechsler scales reported in Table 1. Of the four tests from Table 1 which showed validity coefficients (in terms of Wechsler scores) higher than those obtained with the Shipley-Hartford in the present samples, two involve the expenditure of some professional time, while the other two (AGCT and seven test Kent battery) require relatively long administration times.

Table 2. Summary Data from the Two VA Psychiatric Samples

Factors	Ranges		Means		S.D.		
	VA MHC (N = 99)	VAH (N = 251)	VA MHC	VAH	VA MHC	VA	*t*
Age	17–59	21–72	30.31	35.54	7.53	10.40	5.201**
S–H raw score							
Vocab.	9–40	7–40	27.72	25.79	5.96	7.10	
Abstr.	0–38	0–40	20.28	16.64	10.76	10.04	
Total	17–74	12–77	47.98	42.43	14.80	15.29	3.129*
W–B IQ							
Verbal	73–133	61–136	106.27	103.69	12.35	14.55	
Perf.	67–131	61–136	107.47	105.76	12.08	13.62	
Full sc.	77–131	58–138	107.56	105.28	11.55	13.90	1.561

*Significant at .01 level.
**Significant at .001 level.

TABLE 3. SUMMARY DATA FROM THE STATE HOSPITAL PSYCHIATRIC AIDE SAMPLE

Factors	Ranges		Means		S.D.		
	Male (N = 20)	Female (N = 40)	Male	Female	Male	Female	*t*
Age	23–62	22–64	41.85	46.48	12.97	10.84	1.345
S-H							
Vocab.	16–36	14–39	26.35	28.93	5.17	5.34	
Abstr.	8–32	0–38	18.30	24.55	8.01	9.32	
Total	28–68	21–73	44.65	53.48	11.36	12.92	2.656*
AGCT							
Raw score	51–126	44–119	85.10	89.53	19.15	21.15	
Score	80–141	75–136	108.10	111.68	15.59	17.10	.795

*Significant at the .05 level.

The data in Table 4 indicate that in both VA samples, Shipley scores are more closely related to W-B *weighted* scores than to IQs. This is probably due to the age correction on the Wechsler, a fact which is reflected in the less than perfect correlation between W-B weighted scores and IQs. (In the VAH sample this value was .960 and in the MHC sample .965.) Also in both samples, the best predictor of Wechsler full scale weighted score was Shipley total raw score. A regression equation for predicting Wechsler full scale *weighted* scores from Shipley-Hartford total raw scores was derived for each of the two VA samples. The equation from the VA MHC data was 103.74+.9166 (Shipley total raw score −47.98) and from the VAH data it was 97.33 + 1.1363 (Shipley total raw score −42.43). A table for the conversion of Shipley-Hartford total raw scores to Wechsler *total weighted score* estimates was constructed and is reported as Table 5. It should be noted that in order to obtain

TABLE 4. PEARSON PRODUCT-MOMENT COEFFICIENTS OBTAINED BETWEEN SHIPLEY-HARTFORD RAW SCORES AND SCORES ON THE WECHSLER-BELLEVUE SCALES AND AGCT

Sample	S-H (raw scores)			SEest.
	Vocab.	Abstr.	Total	
VA MHC (N = 99)				
W-B IQ				
Verbal	.762	.544	.703	
Performance	.517	.506	.590	
Full scale	.732	.586	.720	8.015
W-B weighted score				
Verbal	.761	.554	.718	
Performance	.474	.585	.624	
Full scale	.709	.647	.766	11.373
VAH (N = 251)				
W-B IQ				
Verbal	.743	.682	.793	
Performance	.526	.539	.598	
Full scale	.694	.671	.765	8.953
W-B weighted score				
Verbal	.751	.718	.821	
Performance	.498	.613	.634	
Full scale	.676	.735	.793	13.362
State hosp. psych. aides (N = 60)				
AGCT raw score				
Male	.724	.626	.771	
Female	.814	.644	.801	
Total	.785	.635	.780	8.081

the estimated Wechsler full scale IQ the estimated full scale *weighted* score obtained from Table 5 must be referred to the appropriate age column in the tables provided by Wechsler[23].

A similar treatment of Shipley-Hartford and AGCT scores in the state hospital psychiatric aide sample yielded figures comparable with those reported by Manson and Grayson[14], and are therefore not reproduced here. Note that the score equivalents of Shipley-Hartford performance reported by the above investigators (which are used in some clinical settings) are AGCT scores rather than Wechsler-type IQs. Because of the greater variance of AGCT scores than of W-B IQs, the former cannot be considered numerically equivalent to Wechsler IQs, as AGCT scores tend to be several points higher for scores above the mean and lower for scores below the mean.

TABLE 5. WECHSLER-BELLEVUE FULL SCALE WEIGHTED SCORE ESTIMATES FROM SHIPLEY-HARTFORD TOTAL RAW SCORES ON TWO VA PSYCHIATRIC PATIENT SAMPLES

S-H total raw score	W-B wtd. MHC	F. S. sc.* VAH	S-H total raw score	W-B wtd. MHC	F. S. sc. VAH	S-H total raw score	W-B wtd. MHC	F. S. sc. VAH	S-H total raw score	W-B wtd. MHC	F. S. sc. VAH
80	133	140	62	117	120	44	100	99	26	84	79
79	132	139	61	116	118	43	99	98	25	83	78
78	131	138	60	115	117	42	98	97	24	82	76
77	130	137	59	114	116	41	97	96	23	81	75
76	129	136	58	113	115	40	96	95	22	80	74
75	128	134	57	112	114	39	95	93	21	79	73
74	128	133	56	111	113	38	95	92	20	78	72
73	127	132	55	110	112	37	94	91	19	77	71
72	126	131	54	109	110	36	93	90	18	76	70
71	125	130	53	108	109	35	92	89	17	75	68
70	124	129	52	107	108	34	91	88	16	74	67
69	123	128	51	106	107	33	90	87	15	73	66
68	122	126	50	106	106	32	89	85	14	73	65
67	121	125	49	105	105	31	88	84	13	72	64
66	120	124	48	104	104	30	87	83	12	71	63
65	119	123	47	103	103	29	86	82	11	70	62
64	118	122	46	102	101	28	85	81	10	69	60
63	117	121	45	101	100	27	85	80			

*Refer to appropriate age column in Wechsler tables to obtain IQ.

Table 5 shows that even very low Shipley-Hartford scores yield relatively high Wechsler estimates. There are two main considerations here. First, the nature of the populations from which the samples were drawn is such that extremely low intelligence is almost automatically ruled out. Because mental defectives are uncommon in the VA population (relative to their representation in the general population) extremely low IQs would be predicted very seldom—even on the basis of "chance" scores on objective intellectual measures. Secondly, and in accordance with Garfields's findings[7], low Shipley-Hartford scores (*i.e.*, MAs) may be associated with higher level Wechsler performance than might be anticipated on the basis of data reported by previous investigators. From this fact, and from the apparent low "ceiling" of the test (see Table 5), it appears that the Shipley-Hartford as a measure of intelligence has a rather restricted range. That is, in order to perform above a "chance" level on the test, a fair level of intelligence (at least dull normal) is required while maximum scores in the test are attainable by individuals falling within the "superior", or certainly within even the lower range of the "very superior", category in terms of Wechsler scores.

When Shipley-Hartford scores were converted to Wechsler estimates via the above procedures for the purpose of classifying subjects according to the categories suggested by Wechsler[23], somewhat over one-half of the patients were correctly classified. This figure should be compared with that which would obtain by treating the Shipley-Hartford total score as an MA equivalent for the purpose of intellectual

classification. The percentage of correct classifications by the latter method was only 34.4% in the MHC sample, and 36.3% in the VAH sample. Clearly, the conversion of Shipley-Hartford scores to Wechsler estimates by the method outlined yields a much more accurate classification of subjects. Further, though the absolute difference between actual "hit rate" (54.5% in the MHC sample, and 57.4% in the VAH sample) and "chance" expectancy was small, the difference was statistically significant. "Chance" expectancy was construed as the proportion of cases within the modal category in each sample, *i.e.*, 50.5% in the MHC and 51.4% in the VAH sample falling within the average range of intelligence (IQ 90-110).

As noted in Table 4, the standard error of estimate in predicting Wechsler total weighted scores from Shipley-Hartford total raw scores, was 11.37 in the MHC sample, and 13.36 in the VAH sample. As a further evaluation of the method, the mean discrepancy between predicted and obtained IQ, after converting predicted full scale weighted score to IQ in terms of the patient's age, was computed and found to be 6.2 points in the MHC sample, and 6.8 points in the VAH sample.

Summary

The literature contains a number of studies comparing the results of brief measures of intellectual functioning to standard criterion measures such as the Wechsler-Bellevue scales. Most of the more popular measures correlate in the vicinity of .70 with the Wechsler though some, notably those either requiring some professional time expenditure in administration or the longer paper-and-pencil tests, yield somewhat higher validity coeficients. Data from two VA psychiatric samples, and from a sample of state hospital psychiatric aides indicated that the Shipley-Hartford Retreat Scale is as good a measure of current functioning intelligence as are most other brief group measures of intelligence used in clinical practice. Correlations between Shipley total raw scores and the standard intelligence tests used were .77, .79, and .78 with W-B in the MHC, VAH, and with the AGCT in the psychiatric aide samples, respectively. Shipley total raw scores yielded the best estimate of Wechsler scores, and were more highly related to Wechsler total *weighted scores* than to IQs. The standard error of estimate in predicting Wechsler total weighted scores from Shipley total raw scores was 11.37 in the MHC sample and 13.36 in the VAH sample. The comparable figure in predicting AGCT raw scores was 8.08 in the psychiatric aide sample. The mean error in estimating Wechsler IQ by this technique was 6.2 in the MHC sample and 6.8 in the VAH sample. Regression equations were derived for predicting Wechsler full scale weighted scores from Shipley total raw scores, and conversion tables for determining the former were presented for each of the VA psychiatric samples.

In view of the above, and from the fact that the Shipley-Hartford requires no professional time for administration or scoring, it is concluded that the test is a valuable brief measure of current intellectual functioning in clinical settings similar to those from which the data were obtained.

References

1. Abrams, E. N. Prediction of intelligence from certain Rorschach factors. *J. clin. Psychol.*, 1955, *11*, 81-83.
2. Aita, J. A., Armitage, S. G., Reitan, R. M. and Rabinowitz, A. The use of certain psychological tests in the evaluation of brain injury. *J. gen. Psychol.*, 1947, *37*, 25-44.
3. Allen, R. M., Thornton, T. E. and Stenger, C. A. Ammons and Wechsler test performances of college and psychiatric subjects. *J. clin. Psychol.*, 1954, *10*, 378-381.
4. Corsini, R. The Immediate Test. *J. clin. Psychol.*, 1951, *1*, 127-130.
5. Delp, H. A. Correlations between the Kent EGY and the Wechsler batteries. *J. clin. Psychol.*, 1953, *9*, 73-75.
6. Doppelt, J. E. Estimating the full scale score on the Wechsler Adult Intelligence Scale from scores on four subtests. *J. consult. Psychol.*, 1956, *20*, 63-66.
7. Garfield, S. L. The Shipley-Hartford Retreat Scale as a quick measure of mental status. *J. consult. Psychol.*, 1947, *11*, 148-150.
8. Garfield, S. L. A comparison of the Wechsler-Bellevue and Shipley-Hartford Scales as measures of mental impairment. *J. consult. Psychol.*, 1948, *12*, 259-264.

9. GOLDFARB, W. An investigation of reaction time in older adults, and its relationship to certain observed mental test patterns. *Teach. Coll. Contr. Educ.*, 1941, No. 831.
10. GURVITZ, M. S. The Hillside Short Form of the Wechsler-Bellevue. *J. clin. Psychol.*, 1951, *7*, 131-134.
11. JONES, A. W. and RICH, T. A. The Goodenough Draw-A-Man Test as a measure of intelligence in aged adults. *J. consult. Psychol.*, 1957, *21*, 235-238.
12. KNOTT, J. R., CANNICOTT, R. G., UMBERGER, J. P. and BILODEAU, I. M. Brief tests of intelligence in the psychiatric clinic. *J. clin. Psychol.*, 1951, *7*, 123-126.
13. LEWINSKI, R. J. The Shipley-Hartford Scale as an independent measure of mental ability. *Educ. psychol. Measmt.*, 1946, *6*, 253-259.
14. MANSON, M. P. and GRAYSON, H. M. The Shipley-Hartford Retreat Scale as a measure of intellectual impairment for military prisoners. *J. appl. Psychol.*, 1947, *31*, 67-81.
15. MCNEMAR, Q. On abbreviated Wechsler-Bellevue scales. *J. consult. Psychol.*, 1950, *14*, 79-81.
16. MERRILL, R. M. and HEATHERS, L. B. A comparison of the Wechsler-Bellevue and ACE tests on a university counseling center group. *J. consult. Psychol.*, 1953, *17*, 63-66.
17. RABIN, A. I. The use of the Wechsler-Bellevue Scales with normal and abnormal persons. *Psychol. Bull.*, 1945, *42*, 410-422.
18. RAPAPORT, D., GILL, M. and SCHAFER, R. Diagnostic psychological testing. Chicago: Year Book Publishers, 1945.
19. ROBINOWITZ, R. Performances of hospitalized psychiatric patients on the Kent Emergency Test and the Wechsler-Bellevue Intelligence Scale. *J. clin. Psychol.*, 1956, *12*, 199-200.
20. SHIPLEY, W. C. A self administering scale for measuring intellectual impairment and deterioration. *J. Psychol.*, 1940, *46*, 336-377.
21. STACEY, C. L. and GILL, M. R. The relationship between Raven's Colored Matrices and two tests of general intelligence for 172 subnormal adult subjects. *J. clin. Psychol.*, 1954, *11*, 86-87.
22. TAMMINEN, A. W. A comparison of the Army General Classification Test and the Wechsler-Bellevue Intelligence Scales. *Educ. psychol. Measmt.*, 1951, *11*, 646-655.
23. WECHSLER, D. *The measurement of adult intelligence.* (3rd ed.) Baltimore: Williams and Wilkins, 1944.
24. WRIGHT, M. E. Use of the Shipley-Hartford Test in evaluating intellectual functioning of neuropsychiatric patients. *J. appl. Psychol.*, 1946, *30*, 45-50.

A PROPOSED METHOD FOR MEASURING POTENTIAL INTELLIGENCE

ALVIN R. MAHRER AND LEWIS BERNSTEIN

Veterans Administration Hospital, Denver, Colorado[1, 2]

PROBLEM

In the literature on intelligence testing, there is frequent reference to a distinction between measured intelligence and some higher level of potential intelligence which is regarded as unmeasurable. From Rotter's construct of behavior potential (3), it is hypothesized that an individual has had a large body of experiences that are directly or indirectly related to any intelligence test question. From these experiences one may have developed a group of responses to any single test item. Each test question is presumed to serve as a cue for a group of answers.

For example, one Wechsler Bellevue item asks: "What should you do if while sitting in a theatre you were the first person to see smoke and fire?" For a particular individual one set of answers might be the following: "I don't know," "I wouldn't get panicky," "I'd call the fire department," "I'd tell the manager," and "I'd probably get out fast." Each of these answers, or behaviors, is presumed to have a certain *behavior potential* relative to the other answers. The behaviors may then be described as falling along a continuum of decreasing behavior potential. Those with the higher behavior potential will be given as a response more readily than those with lower behavior potential, whereas standard testing procedure assumes that the answer elicited by the test item is the best response of which the individual is capable.

For example, if a hospital patient sees the testing situation as one in which he is being clinically evaluated, and if he has a high expectancy that by not revealing himself, and by avoiding "pathological" answers he might soon be able to leave the hospital, then one might predict that

---

[1]The data for this study were collected while the senior author was stationed at Fitzsimons Army Hospital, Denver, Colorado.

[2]The authors wish to express their appreciation to Dr. Victor C. Raimy for his critical reading of the manuscript.

the behavior potential of avoidant responses such as "I wouldn't get panicky," or "I don't know," would perhaps be increased beyond that of "I'd tell the manager." The result would be that he would fail the item even though a better answer is among his group of potential behaviors.

To summarize, the usual method of intelligence testing elicits *one* response for each question. The usual intelligence test score is, therefore, a product of only those responses with the highest behavior potential in a particular situation. Those responses given in a specific situation may not be the "most correct" responses in the patient's repertory. However, if the intelligence tester asks for a *group* of responses to each test question, the scoring could be based on the best answer among those given by the subject. Such a method of intelligence test administration and scoring provides an operational definition of potential intelligence. It is the purpose of this study to test this concept of potential intelligence using the proposed method of measurement. Specifically, it is hypothesized that the potential intelligence score will be significantly higher than the standard intelligence score.

Method

Subjects were 120 consecutive patients referred for testing to the Clinical Psychology Service at Fitzsimons Army Hospital, exclusive of military prisoners, children, and emergency cases. There were 110 male and 10 female patients. Ninety subjects were military personnel and 30 were dependents of military personnel. Sixteen subjects were out-patients and 104 were in-patients. The mean age of the experimental group was 24.18 years, with a SD of 5.90; for the control group the mean age was 27.93 years, with a SD of 8.68.

The test used was the Wechsler-Bellevue Intelligence Scale, Form I[4]. The scores on several of the Wechsler-Bellevue subtests depend, in part, upon speed of response, immediate memory, or the grasping of a general principle. These subtests are not appropriate to the repeated testing required in this study. Accordingly, the following five verbal subtests were used: Information, Comprehension, Arithmetic, Similarities, and Vocabulary.

All subjects were first tested with the standard administration, except that *all* items of the five subtests were included. On the basis of the IQ score on this administration (derived from the five subtests used in this study), each subject was then assigned to the experimental or the control group according to a matching procedure (N = 60 in each group). Immediately following the initial testing, a second standard administration of the test followed for both groups. The second IQ was determined by scoring the better of the two responses to each item from each of the first two administrations of the test.

Following the second administration, the control group was tested a third time, with the following instructions: "Now we are going over the questions once more. Remember to give the best answer you can think of. This is not a test of memory, so you may give new answers if you think they are better." The third IQ for the control group was determined by scoring the best of the three responses to each item from the three administrations of the test.

The third administration of the test to the experimental group was designed to elicit responses (behavior potentials) in addition to those given in the standard administration. The following instructions were given to this group: "I am interested in the way people guess. Since you can't guess when you already know the right answer, I am going to go back over the questions you didn't get completely right. When I ask you the question you are to give four answers, whether you think the answers are right or wrong. The important thing is to give four answers to each question. You may guess as much as you want."

If these instructions failed to elicit four answers, the following instructions were used, in sequence: (a) "What is your 2nd (3rd, 4th, or next) answer?" (b) "What is the next guess you think of?" (c) "Try and make a wild guess." (d) "Tell me anything you think of when I ask this question." (e) "If you can't think of any answer, then make up an answer."

The third, fourth, fifth, and sixth IQs for the experimental group were determined by scoring the item with the highest score as each additional response to the third administration was considered in the scoring. For example, the fifth IQ score was based upon the best response from the first two administrations and the first three responses to the third administration. It should be noted that by using this method of scoring, each succeeding IQ must be either equal to or higher than the preceding IQ.

Because of the premium of speed of response in Arithmetic items 9 and 10, these items were not included in any of the retests for either group when the subject achieved a score greater than zero. Similarly, the score on the fifth Comprehension item depends on the number of elements of information; therefore, this item, too, was excluded from further retesting once a score greater than zero was obtained. With these exceptions, scoring followed the standard procedure.

Results

Table 1 indicates that, when the mean IQ scores of the experimental and control groups are compared, there are no significant differences between the groups follow-

Table 1. Mean IQ Scores of Experimental and Control Groups

Group		*N*	Test Administration 1	2	3*	3a	3b	3c
Exper.	(Total)	60	100.56	101.04	110.25	112.83	115.20	117.63
Control	(Total)	60	99.43	100.82	101.06			
Exper.	(Upper 1/3)	20	116.40	116.72	125.33	127.52	128.20	131.05
Control	(Upper 1/3)	20	116.53	118.13	118.80			
Exper.	(Middle 1/3)	20	101.15	101.39	110.09	113.09	118.31	120.20
Control	(Middle 1/3)	20	99.60	101.04	101.05			
Exper.	(Lower 1/3)	20	84.15	85.01	95.33	97.88	99.09	101.65
Control	(Lower 1/3)	20	81.85	83.29	83.35			

*For the experimental group, test administration 3 refers to the first of the four responses elicited in the third administration; 3a refers to the first and second of the four responses; etc.

In comparing the experimental and control groups under administrations 1 and 2, none of the *t*-values are significant. Under administration 3, all *t*-values are significant at or beyond the .01 level of confidence.

ing the second administrations of the test, indicating that simple readministration does not appreciably change the IQ level. When the test is readministered a third time to the control group, there is still no appreciable IQ change. However, the third administration for the experimental group (where four responses to each test item were elicited) produced a significant rise in IQ. This finding is interpreted as supporting the hypothesis that the potential intelligence score would be significantly higher than the standard intelligence score. When the subjects in each group are divided into upper-, middle-, and lower-thirds of IQ scores, it is apparent that the potential intelligence scores are significantly higher throughout the range of IQs represented in this study (See Table 1). It might be expected that the lower the initial IQ, the greater would be the difference between potential and standard IQ. However, it is of some interest to note that the amount of IQ increase for each one-third of IQ scores (administration 3c minus administration 1, Table 1) is approximately the same for each one-third of standard IQs.

In the experimental group, the mean increase of potential IQ over standard IQ (administration 3c minus administration 1, Table 1) was 17.25 points. The S.D. was 5.37 points. The range was from 7.0 to 29.5 points. Four cases increased from 5 - 9.9 IQ points; 15 cases increased from 10 - 14.9 points; 24 cases from 15 - 19.9 points; nine cases from 20 - 24.9 points; and eight cases increased 25 - 29.9 points. These data point to the existence of individual differences in the increases of potential over standard IQ scores.

In comparing the experimental with the control group following the third administration of the test, we can justifiably use only the first of the four responses elicited from the experimental group, as the control group gave only three responses. However, it is of some interest to note that, as each of the additional responses given by the experimental group is used in computing the mean IQs, the potential IQ scores become progressively larger (See Table 1). This fact adds additional support to our hypothesis that, as a *group* of behavior potentials is elicited, the potential intelligence score will be significantly higher than the standard intelligence score.

Examination of Table 2 indicates the potential IQ score is independent of the diagnostic categories represented in this study. A criterion *t*-test was computed between the "other psychoses" and "neurotic" groups, since this difference was the largest. The *t*-value of .78 was not significant.

TABLE 2. MEAN STANDARD AND POTENTIAL IQ SCORES OF EXPERIMENTAL GROUP BY DIAGNOSTIC CATEGORY

Diagnosis	N	Standard IQ	Potential IQ	Increase
Schizophrenia	17	101.05	117.70	16.65
Schizoid & Paranoid Personality	8	105.75	124.37	18.62
Other Psychoses	5	91.20	111.20	20.00
Character Disorder	15	93.66	110.73	17.07
Organic	4	90.25	107.50	16.25
No disease	4	107.75	127.50	19.75
Neurotic	7	111.00	125.57	14.57

The data in Table 3 indicate that none of the five subtests contributes more to the increase in potential IQ than any other subtest. A criterion *t*-test was calculated between the subtest with the largest increase (Arithmetic) and the subtest with the smallest increase (Vocabulary). The *t*-value of .76 was not significant.

TABLE 3. MEAN STANDARD AND POTENTIAL WEIGHTED SCORES OF THE EXPERIMENTAL GROUP BY SUBTEST

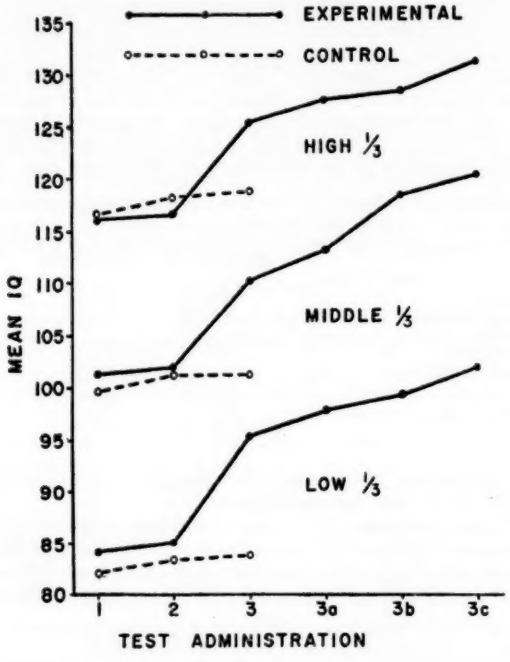
Subtest	Standard Weighted Score	Potential Weighted Score	Increase
Information	9.40	11.58	2.18
Comprehension	9.93	13.45	3.52
Arithmetic	8.33	11.75	3.42
Similarities	10.22	13.85	3.63
Vocabulary	8.95	10.68	1.73

DISCUSSION

The results of this attempt to measure potential intelligence appear promising. However, certain deficiencies of this study should be pointed out. It is apparent that the third administrations of the test to the control and experimental groups were not strictly comparable. For the control group, a complete third administration of the test was conducted, whereas the experimental group was instructed that they were being retested only over those items which they did not get completely right. It may be, therefore, that the increase in potential IQ scores was, at least in part, a function of the differential feedback to the two groups. It would appear that in any attempt to repeat this study, the control group should be given a similar kind of feedback.

It was also observed that it was progressively more difficult to elicit answers beyond the first response to each question. More experimentation needs to be done with different ways of economically eliciting responses for various groups. For example, would children be as reluctant to give a group of answers as were the adult military psychiatric patients in this study? Could our instructions be altered to elicit responses more readily?

FIGURE 1. MEAN IQ EXPERIMENTAL CONTROL GROUPS DIVIDED INTO UPPER, MIDDLE AND LOWER ONE-THIRDS OF IQ SCORES.



From Figure 1, it is apparent that as each additional response from third administration to the experimental group is considered in the scoring, the mean IQ score rises. The curves do not appear to level off. If we had been able to elicit five or six (or more) responses to the final administration, would the curves have continued to rise? Further work will need to be done to determine at what point a leveling-off will take place.

If the findings in this study can be substantiated in further work with a variety of groups, certain clinical implications would follow. For example, we might make different clinical decisions about a patient whose standard IQ was in the mental defective range, but whose potential intelligence was within the dull normal or normal range, as compared with a person whose standard and potential IQs were the same[1]. Our technique might also be valuable in assessing the amount of intellectual loss from organic brain damage, and for predicting the extent of probable recovery of intellectual function following retraining.[3]

SUMMARY

Rotter's construct of behavior potential suggests that an individual has a large body of experiences that are directly or indirectly related to any intelligence test item. Those experiences with the higher behavior potential will be given as a response to an intelligence test question more readily than those with lower behavior potential. Since the usual intelligence test elicits only one response to each item, the IQ is a product of only those responses with the highest behavior potential in a particular situation. In the present study, an attempt was made to elicit a group of

[3]For a general discussion of clinical implications of potential intelligence, see MAHRER[1].

responses to each test item, with the IQ based upon the best answer from the group. Such a method of administration and scoring provided an operational definition of potential intelligence. Specifically, it was hypothesized that the potential intelligence score would be significantly higher than the standard intelligence score.

Our findings supported the general hypothesis. It was also found that the potential IQ increased throughout the range of IQs represented in this study. The amount of increase was not significantly different between the upper-, middle-, and lower-third of standard IQs. Potential IQ was found to be independent of psychiatric diagnosis. No one subtest contributed more to the potential IQ than any other subtest. Methods of improving the measurement of potential intelligence were discussed, and several clinical implications of our findings were suggested.

References

1. Mahrer, A. R. Potential intelligence testing: a case study. *U. S. Armed Forces med. J.*, 1957, *8*, 684-692.
2. Mahrer, A. R. Potential intelligence: a learning theory approach to description and clinical implication. *J. gen. Psychol.*, In press.
3. Rotter, J. B. *Social learning and clinical psychology.* New York: Prentice-Hall, 1954.
4. Wechsler, D. *The measurement of adult intelligence.* (3rd ed.) Baltimore: Williams & Wilkins, 1944.

INFERENCE OF INTELLECTUAL EFFICIENCY FROM THE WAIS VOCABULARY SUBTEST

STEPHEN L. FINK AND FRANKLIN C. SHONTZ

Highland View Hospital, Cleveland 22, Ohio

Problem

In clinical practice one of the most important uses of the Wechsler intelligence scales is to provide not only a quantitative estimate of the subject's intellectual potential but a qualitative evaluation of his functioning.

One important qualitative aspect of intellectual behavior is "efficiency", operationally defined to exist when, on a series of items of graded difficulty, the subject passes all items up to the level of his maximum capacity and then fails all items beyond that point. Schafer comments that in test records of neurasthenics and schizophrenics, "Temporary inefficiencies are usually most striking: relatively easy items of Information, Vocabulary, Digit Span, Arithmetic, Picture Completion, and Block Designs tend to be failed, while more difficult items are passed" (1, p. 40; *cf* also pp. 64-65).

The quotation cited refers to the Wechsler-Bellevue examination, but it may be reasonably extended to refer to the WAIS (2) as well. The statement and the definition derived from it assume "efficiency" to be the normal characteristic of a healthy person's intelligence test record; however, neither take into account the possibility that certain indicators of inefficiency may be "built into" the test and are not to be considered valid qualitative indices of poor intellectual functioning.

In the present study, the Vocabulary subtest of the WAIS was examined to determine the degree to which "inefficiency" may be inferred from the type of signs usually employed. No effort was made to test the predictive validity of the concept of intellectual efficiency. The sole purpose was to examine some of the conditions under which inefficiency may be logically inferred on the basis of its operational definition. The Vocabulary subtest of the WAIS was selected because the function it measures holds up well with age, is highly reliable, is probably the best single index of verbal intelligence (2, 3), and is constructed in such a way as to suit it well to quantitative analysis.

Method

Subjects. One hundred protocols from Wechsler's original standardizing group were obtained by special arrangement with the Psychological Corporation. The sample was selected to be representative of the overall standardizing population of the WAIS. Vocabulary subtest weighted scores for this group (according to Wechsler's scorings) were normally distributed, with a mean of 10.2 points (SD = 3.0).

A *validation* group of one hundred patients hospitalized with chronic physical illnesses was also examined to provide a cross-check upon the findings from the original sample. The mean age of this group was 41 years (SD = 15.2). The distribution of weighted scores was normal with a mean of 9.2 points (SD = 3.5). The validating group therefore included a much larger proportion of aged individuals than did Wechsler's group, and the number of male subjects (N = 56) was slightly greater than the number of female subjects (N = 44). Furthermore, the mean weighted score was significantly lower (p .05) in the validation sample than in Wechsler's sample. Although the hospitalized group could not be considered to be a representative sample of the general population, it was felt that any trends which were discovered in Wechsler's group and confirmed in this group could be considered reliable, since they obviously would have to occur in spite of variations in the validation sample from general population representativeness. It is worthwhile to note that the distribution of Vocabulary scores for the validation group differed only slightly from that obtained from Wechsler's original subjects.

Procedure. The WAIS Vocabulary subtest was administered to the validation group and scored in the standard manner. The only deviation from standard procedure was that the first three items (*bed, ship, penny*) were administered to subjects in this group as a check upon their original placement. Scorings provided by Wechsler for the sample from the standardizing population were accepted without alteration.

For each group, a table was constructed showing the frequency of occurrence of zero-point scores for every item. Similar tables were then constructed for one and two-point scores. The method of least squares was used to establish expected frequencies for each of the six tables. Where this technique failed to provide a straight-line equation which met the requirements of a broad-category test of "goodness-of-fit", combined graphic and least squares techniques were used to isolate those stimuli which were unduly influencing the estimates. In these cases, a method of successive elimination of deviant items and successive recalculation of expected frequencies finally yielded sets of data which met the requirements of a rigid test of "goodness-of-fit". A comparison was then made of items eliminated by this technique in both samples of data.[1]

All deviations from expected frequencies were evaluated by chi-square, and a level of significance of .001 (two-tailed) was established as critical for rejection of the null hypothesis.

Results

Obtained frequencies of zero-point scores showed almost perfectly linear increases from the beginning to the end of the subtest in both the Wechsler and the validating sample of protocols. Generally linear trends were also apparent in the frequency of occurrence of two-point scores, except, of course, that there was a progressive decrease from the beginning to the end of the subtest. In addition, certain items displayed frequency values which were far out of line with expectations. In every case, the deviation was in the direction of there being fewer two-point scores than would be expected, and in every case, deviations were consistent in both samples (Table 1). The wide disparity between the observed values in the two samples for the item *slice* is due to the fact that the present authors were far more strict in applying a specific scoring criterion (*i.e.*, that two points be given only for the idea of "to cut *thinly*") than were Wechsler's scorers. In 54 cases from the Wechsler sample

TABLE 1. ITEMS SHOWING SIGNIFICANT DEVIATIONS FROM EXPECTED FREQUENCES OF TWO-POINT SCORES*

Item	Observed frequency (Wechsler's sample)	Expected frequency (Wechsler's sample)	Observed frequency (Validating sample)	Expected frequency (Validating sample)
Winter	38	82	49	72
Slice	63	73	5	64
Enormous	33	67	27	58
Sentence	26	63	30	55
Regulate	44	61	33	53
Remorse	19	42	10	35

*Multiplicatively combined probabilities for both independent samples less than .001.

the response to "slice" did not include the idea of thinness but was given a two-point scoring anyway.

Analysis of the frequency of one-point scores was made difficult by the fact that straight lines could not be found by the least squares method, which would satisfy even gross tests of "goodness of fit". However, since this method consistently established lines nearly parallel to the abscissa (item numbers), items with frequencies which grossly deviated from such a general line were successively eliminated from the distribution. Eliminations of this type finally yielded lines which satisfied the requirements of a rigid chi-square test of "goodness-of-fit". No words were eliminated from the Wechsler sample which were not also eliminated from the validating sample, although a few words were found to be deviant in the validating group but not in the Wechsler group. Words which it was necessary to eliminate in both samples are presented in Table 2.

TABLE 2. ITEMS ELIMINATED FROM BOTH SAMPLES DUE TO SIGNIFICANT DEVIATIONS IN FREQUENCY OF ONE-POINT RESPONSES

Item Number	Item
1	Bed*
2	Ship*
3	Penny*
4	Winter
6	Breakfast
7	Fabric
8	Slice
11	Enormous
13	Sentence
14	Regulate
23	Remorse
29	Tranquil*

*Eliminated due to significantly low frequencies of one-point responses. All others eliminated due to significantly high frequencies of one-point responses.

DISCUSSION

It would appear that Wechsler arrived at his original item placements from a consideration of the frequency distribution of zero-point scores, since these were almost perfectly linear in both samples. The present study therefore confirms the item placements as they appear in the published scale.

It is evident that at least six items of the Vocabulary subtest are scored with a value of two points a significantly low proportion of the time. The items are: *winter, slice, enormous, sentence, regulate,* and *remorse.* These six items were also found to be scored with a value of one point a significantly high proportion of the time. It is safe to propose, therefore, that one-point responses for these items have a strong "pull"

and should not be considered as signs of intellectual inefficiency even if they appear to be out of place in an individual record.

The words *breakfast* and *fabric* (the latter of which was actually of borderline significance in the analysis of two-point responses) also "pull" toward one-point scores, although they do not seem to do so at the expense of other scoring values.

The word *tranquil* tends to pull away from a one-point scoring value. Since it is not out of line with respect to the other scoring categories, however, other responses are probably equally distributed between zeros and twos. This would imply that the item is either passed or failed, but is not often given a partial score. Consequently, it would seem quite possible for a subject to obtain an "out of place" zero or two on this item without indicating any inefficiency of performance with respect to the rest of the scale.

The pull of the items *bed*, *ship*, and *penny* is, of course, easily explained by the fact that one-point responses are not possible on these items.

It is suggested that the clinician take into account the findings of the present research if he wishes to use the concept of *intellectual efficiency* in his analysis of WAIS Vocabulary protocols. The most conservative way to accomplish this would be to eliminate from consideration all the items which have been brought into question here and to consider all one-point responses as "passes" rather than as "partial passes" or complete failures. Item pull will affect the clinician's judgment considerably less if this procedure is followed, since the subject will always be given the benefit of comparison against the most stable standards possible.

Summary

Frequency of zero-, one-, and two-point responses were analyzed in two independent samples of 100 WAIS Vocabulary subtest protocols. One sample of protocols was made up of a systematically randomized selection from Wechsler's original standardizing group; the other was obtained from 100 patients with chronic illnesses. The stimulus words *winter*, *slice*, *enormous*, *sentence*, *regulate*, and *remorse* were found to obtain a significantly low frequency of two-point scores and a significantly high frequency of one-point scores. The items *breakfast* and *fabric* were found to obtain a significantly high frequency of one-point scores, but not at the expense of the other two scoring categories. The single item *tranquil* was found to obtain a significantly low frequency of one-point scores. Implications of the findings for qualitative analysis of WAIS Vocabulary protocols are discussed.

References

1. Schafer, R. *The clinical application of psychological tests.* New York: International Universities Press, 1948.
2. Wechsler, D. *The measurement of adult intelligence.* Baltimore: Williams & Wilkins, 1944.
3. Wechsler, D. *Manual for the Wechsler Adult Intelligence Scale.* New York: Psychological Corp., 1955.

RELATIONSHIPS BETWEEN THE BENDER GESTALT TEST AND THE WECHSLER INTELLIGENCE TEST FOR CHILDREN

ELIZABETH MUNSTERBERG KOPPITZ

Children's Mental Health Center, Columbus, Ohio

This study attempts to discover the relationship between performance on the Bender Gestalt Test[1] and intelligence as measured on the WISC. It has been shown that the Bender is a useful tool in diagnosing learning problems in young school children[2, 3]; however, it is also well known that brighter children tend to be better students than duller children. Therefore, the question presents itself whether performance on the Bender is primarily a function of intelligence, or whether it depends on functions different from those measured by the WISC[4].

METHOD

Subjects. Ninety elementary school children, including 20 first graders, 20 second graders, 25 third graders, and 25 fourth graders, served as Ss for this study. They had an age range from 6 years, 7 months to 11 years, 7 months with a mean age of 8 years, 8 months. Their mean WISC IQ was 95, range from 73 to 126. All Ss were clinic or private patients who had been referred either primarily because of behavior problems or because of learning problems. All were given the WISC and the Bender Gestalt Test in one session at the time of psychological evaluation by the author or by one other psychologist.*

Procedure. All WISC protocols were scored according to standard procedure. The Bender protocols were scored according to a scoring system developed and validated by the author[3] for the use with young school children. This scoring system includes four major scoring categories with 31 separate items listed in abbreviated form below:

I. *Distortion of Shape.* Fig. A: a. square or circle grossly misshapen; b. square and circle grossly disproportioned; c. sides of square do not meet.
Fig. 1, 3, 5: two or more dots converted into circles or dashes.
Fig. 6: a. more than three angles in curves; b. no curves, straight line.
Fig. 7: hexagons grossly disproportioned.
Fig. 7, 8: angles extra or missing.

II. *Rotation.* Fig. A, 1, 2, 3, 4, 5, 7, 8: rotation of design or part of it by 45° or more.

III. *Failure of Integration.* Fig. A, 4, 7: failure to join parts; or excessive overlap.
Fig. 2: one or two rows of dots omitted; four or more dots in most columns.
Fig. 3: a. rows of dots added or omitted; shape lost. b. lines instead of dots.
Fig. 5 a. five or less dots in arch; shape lost. b. lines instead of dots.
Fig. 6: two lines not crossing or crossing at extreme end of one line.

IV. *Perseveration.* Fig. 1: more than 14 dots.
Fig. 2: more than 14 columns of circles.
Fig. 6: more than 5 sinusoidal curves in either direction.

In this scoring system, deviations from the stimulus cards are scored. Therefore a high Composite Score on the Bender indicates a poor test performance, while a low Composite Score indicates a good Bender. All Bender protocols were scored by the writer blindly, that is, without knowledge of the Ss' age, grade and IQ.

Since the WISC is scored by comparing each child's performance with that of other children of the same age, the Bender scores were also evaluated on the basis of the age norms for each child. The Mean Composite Scores on the Bender were previously obtained from several hundred children tested in the child guidance clinic and in schools. They are as follows:

*The writer wants to thank David D. Blyth for his valuable help.

Age 6	Mean Composite Score on Bender					10
Age 7	"	"	"	"	"	7
Age 8	"	"	"	"	"	6
Age 9	"	"	"	"	"	5
Age 10	"	"	"	"	"	4

In this study a child was considered to have a good Bender protocol if his Composite Score was at or below the Mean Composite Score for his age level.

The relationship between the WISC and the Bender was determined by computing four-fold Chi-squares for the first and second graders, that is, the number of Ss with good Benders and below or above average IQ's was compared with the number of Ss with poor Benders who had above or below average IQ's. The same procedure was followed for the third and fourth graders, and for all four grades combined. In this way the Bender scores were compared with the WISC FSIQ, VIQ, PIQ, and with each of nine WISC Subtests (I, C, A, S, PC, PA, BD, OA, Cod.).

Thereafter all Ss were divided into those who had been referred primarily because of behavior problems (N = 28) and those with learning problems (N = 62). These two groups were then compared on their WISC IQ's and Bender Composite Scores. Chi-Squares were then computed for the first two grades, the third and fourth grades, and for all four grades together, to determine which of these measures was most closely related to learning prodlems.

Results and Discussion

Table 1 shows the statistical findings when the various WISC measures were compared with the Bender Composite Scores by means of Chi-Square tests.

Table 1. A Comparison of Bender and WISC Scores for Young School Children

WISC	Bender-Gestalt Composite Scores					
	1 & 2 Graders (N 40)		3 & 4 Graders (N 50)		All 4 Grades (N 90)	
	X^2	P	X^2	P	X^2	P
FSIQ	3.0	< .10	4.4	< .05	7.3	< .01
VIQ	5.9	< .02	2.1	> .10	7.4	< .01
PIQ	5.3	< .02	8.1	< .01 > .001	11.9	< .001
I, C, S	not signif.		not signif.		not signif.	
A	2.6	> .10	8.3	< .01 > .001	10.3	> .001
PC	4.9	< .05 > .02	2.1	> .10	6.8	< .01
PA	7.4	< .01	2.4	> .10	10.3	> .001
BD	2.6	> .10	3.9	< .05	7.3	< .01
OA	5.3	< .02	6.6	< .01	12.1	< .001
Cod	not signif.		2.8	= .10	not signif.	
	Children with Learning Problems					
FSIQ	4.0	< .05	14.8	< .001	17.5	< .001
VIQ	not signif.		11.3	< .001	11.2	< .001
PIQ	10.9	< .001	15.8	< .001	26.6	< .001
Bender	9.9	< .01 > .001	2.4	> .10	9.9	< .01 > .001

WISC FSIQ. The relationship between the FSIQ for all 90 Ss and the Bender is highly significant. However, when only the first and second grades are considered this relationship changes to little more than a trend in the positive direction. It increases significantly on the third and fourth grade level. This suggests that the Bender may be more closely related to development and maturation in visual motor perception on the six and seven year old level than to general intelligence.

VIQ. The relationship between the VIQ and the Bender for all 90 Ss together is as significant as for the FSIQ, but the situation changes for the separate grade levels. For the first and second graders the relationship is significant. Maturational factors may be involved in the verbal area as well as in visual motor perception. On the third and fourth grade level this relationship is no longer significant and only shows a trend.

PIQ. The relationship between the PIQ and the Bender is consistently significant at all levels for all grades separately or combined.

WISC Subtests. Information, Comprehension, Similarities: these three tests appear to be tests of verbal reasoning, primarily, and show no significant relationship to the Bender for the separate grade levels nor for all grades combined.

Arithmetic. The relationship between Arithmetic and the Bender is highly significant for all Ss combined. For the first and second graders the relationship indicates a strong trend in the positive direction, while it increases significantly for the third and fourth grades. It appears that Arithmetic and the Bender both deal with whole-part relationships and number concepts.

Picture Completion, Picture Arrangement, Block Design, Object Assembly. As would be expected the Bender shows a significant relationship with all four of these performance tests. The most significant single relationship exists between OA and the Bender for all grades separately as well as together. Since OA also shows the highest correlation with the WISC PIQ it may be possible to substitute the Bender for the WISC Performance Scale for screening purposes or when pressed for time in a clinic situation.

The relationship between BD and the Bender for the first and second graders shows only a positive trend which is much lower than expected. It is felt that this is a function of the BD test at the lower age levels; its character changes markedly at the third grade level. A similar situation exists in PA. This test resembles more the OA test on the six and seven year old level than PA on the higher levels. As a result the relationship between the Bender and the PA test for the first and second grades is highly significant, but drops to a trend only for the third and fourth grades.

Coding. The relationship between Coding and the Bender was not significant for the first and second grades, nor for all grades combined. It shows a positive trend for the third and fourth graders. The fact that Coding scores are primarily based on a time factor, whereas the Bender discards time completely may account for some of the lack of interrelatedness.

Bender, WISC and Learning Problems

Relationships between the WISC FSIQ, VIQ, PIQ, the Bender, and learning problems were explored. Table 1 shows the Chi-Square values and levels of significance. When all four grades are considered together, all four measures listed above show significant relationships to learning problems. This situation changes, however, for the first and second grades alone. On this level there is a very significant relationship between learning problems and both the PIQ and the Bender. The FSIQ is less significantly related to learning problems, while the VIQ shows no significant relationship at all on this level. These findings strongly suggest that learning problems in the first two grades may be closely associated with slow development and immaturity in visual motor perception. It appears further that the Bender may be a good predictor of learning problems in the first two years of school.

On the third and fourth grade level, verbal factors and general intelligence seem to gain in importance in school achievement. It seems that by the third grade most children have matured sufficiently in visual motor perception, so that this area no longer presents the major problem in most cases. The relationship between the Bender and learning problems shows a trend in the positive direction but is no longer significant. At this level the WISC FSIQ, VIQ and the PIQ, all show highly significant relations to learning problems.

Summary

This study explores the relationships between the WISC FSIQ, VIQ, PIQ, and nine WISC Subtests with Bender Composite Scores calculated by the Koppitz Scoring system. Ninety elementary school children served as Ss. They were first through fourth grade students and had a Mean IQ of 95. All were clinic or private patients referred for psychological evaluation because of behavior or learning problems.

Statistical analysis revealed a highly significant relationship between the Bender and the WISC PIQ, FSIQ, VIQ, Arithmetic, and all Performance Subtests with the exception of Coding. These relationships show marked changes at different grade levels. Information, Comprehension, and Similarities do not appear to be related to the Bender.

Learning problems in the first and second grades seem to be primarily related to immaturity in visual motor perception as measured on the Bender Gestalt Test and the WISC PIQ; on the third and fourth grade level, verbal factors and general

intelligence, as well as visual motor perception, appear to be closely related to learning problems.

References

1. Bender, Loretta. *A Visual Motor Gestalt Test and its Clinical Use.* New York: American Orthopsychiatric Association, 1938.
2. De Hirsch, Katrina. Tests designed to discover potential reading difficulties at the six-year-old level. *Am. J. Orthopsychiat. 1957, 27, No. 3.*
3. Koppitz, Elizabeth M. The Bender Gestalt Test and Learning Disturbances in Young Children. *J. clin. Psychol.*, 1958, 14.
4. Wechsler, D. *Wechsler Intelligence Scale for Children.* New York: The Psychological Corporation, 1949.

A MULTIPLE CHOICE SUPRAORDINALITY TEST

HENRY N. PETERS

Veterans Administration Hospital, Jefferson Barracks, Mo.

Problem

An hypothesis based on Korzybski's general semantics led the author to explore the relationship between the making of supraordinate associations and social maladjustment[1, 2]. This work supported the conclusion that, with a controlled free association method, maladjusted subjects show a stronger supraordinate tendency than do normal subjects. The free association supraordinality (FAS) test was judged to be unsatisfactory, because (a) it was difficult to get the desired sets (supraordinate and subordinate) across to *S*s with either spoken or printed instructions, (b) there was an element of subjectivity in the scoring, since with a few ambiguous responses *E* had to decide the question of supraordinality, and (c) the method does not lend itself to group testing. These difficulties would be overcome if a multiple choice form of testing supraordinality (MCS) were found to show the same relationship to maladjustment, and thus could be assumed to be measuring the same thing.

Procedure

The MCS test was printed on two sheets of paper. It included brief instructions and 30 sets of words. Each set consisted of one stimulus word and four response words. *S* was instructed to underline one of the response words, the one which seemed to "go with" the stimulus word. The following are two examples of test items: (a) Color: sensation, noise, dampness, green; (b) Fruit: metal, orange, liquid, food. Each set of four response words contained one supraordinate, one subordinate, and two unrelated words. Two forms of this test were made up and given to 450 subjects in several groups, including college students and hospitalized mental patients (for the most part schizophrenics). The method of scoring was, as with the FAS test, the total number of supraordinate responses minus the number of subordinate.

Results

The first difference to be observed between the results of the two methods of measurement is in the relative frequencies of supraordinate and subordinate associations. Whereas with the FAS the majority of *S*s gave more supraordinates, with the MCS method there was a greater tendency to give subordinates. The FAS means for college sophomores and hospital patients were, respectively, $+3.50 \pm 0.48$ (σm) and $+8.65 \pm 0.59$; with the MCS method the corresponding means shown in

Table 1 are minus for both groups. As the figures of this table indicate, this tendency toward subordinate associations is somewhat dependent upon the particular form of the test and the type of S.

TABLE 1. MEAN MULTIPLE CHOICE SUPRAORDINATE SCORES OF FOUR GROUPS OF VARIOUS EDUCATIONAL LEVELS, ON TWO FORMS OF THE TEST

Groups	Form I			Form II		
	N	M	σm	N	M	σm
Hospital Patients	62	−9.03	1.48	37	−3.02	1.72
College Sophomores	236	−3.80	0.87			
Advanced Students	72	−2.82	1.58	30	+4.51	2.33
Psychologist Ph.D.'s	7	+10.57	5.23	6	+5.33	5.48

A second striking difference in the results of the two methods is in the relative scores of college students and patients. Whereas the FAS gave results predicted from the general semantic hypothesis, namely, that the supraordinate scores of the patients would be higher than those of the college students, with the MCS method just the reverse was found. Assuming that the two methods are measuring the same thing, the results are patently contradictory. The one indicates that the patients are more abstract in their associations, the other that the patients are more concrete.

There are two possible interpretations of this difference. One is that there is a basic difference in what the two methods are measuring. The other is that they are measuring essentially the same thing, but that the effect of education is greater with the MCS procedure. There are two lines of evidence supporting the second of these interpretations. (a) There is evidence that educational level does determine the MCS score, to an extent not found with the FAS method. (b) Within groups of homogeneous educational level, the positive relationship between supraordinality and maladjustment seems to hold true.

The groups of *S*s are listed in Table 1 in order of increasing level of education. The mean supraordinality scores for both forms increase in the same order. The patient mean differs reliably from the other means in the same column. The difference between the Ph.D.'s and the other two nonpatient groups are also significant. It is thus likely that with the MCS method a strong positive relation between level of education and choosing supra-ordinates is stronger than any opposing relation between maladjustment and supraordination. Scrutiny of the printed choices in the MCS forms supports this interpretation. Many of the supraordinates, such as, "isosceles", "communication", and "occupation", are obviously words of lower familiarity than the corresponding subordinates.

The correlation between MCS score and two measures of intelligence has been examined, and no consistent evidence of relationship found. Two groups of patients were given the Raven Progressive Matrices test, as well as one form of the MCS. Between the Matrices and Form I of MCS, $r = -0.11$ ($N = 38$); between Matrices and Form II, $r = +0.09$ ($N = 36$). With college students the two Forms of MCS and two forms of a circular pencil maze test[3] were given to six groups (N varying from 11 to 20). Rho varied from −0.55 to +0.45, and approached significance ($p < 0.05$) in only one case.

Form I of the MCS test was given to 43 college students who came to the student counseling service at Washington University for help. The staff of the counseling service judged 16 of these students to be adjusted, the other 27 to be either maladjusted or doubtful. The supraordinality scores for these two groups are, respectively, -6.25 ± 3.20, and $+3.63 \pm 2.42$. Here the difference is in the direction of greater supraordinality for the nonadjusted students, a difference which is significant at less than the .01 level. The difference between the mean for the nonadjusted and for the large group of college sophomores shown in Table 1 is also significant beyond the .01 level.

TABLE 2. MEAN MCS SCORES OF PATIENTS OF TWO LEVELS OF SYMPTOM MORBIDITY, AT 3 SUCCESSIVE TESTINGS

Tests	High Morbidity N	M	σm	Low Morbidity N	M	σm
1st Test (Form I)	14	−6.08	2.35	16	−9.93	3.30
2nd Test (Form II)	11	+2.01	3.66	16	−8.98	2.06
3rd Test (Form I)	10	−4.96	4.26	9	−19.71	1.72

Thirty-two of the patients of the group cited in Table 1 were used in another research project in which their ward behavior was rated by aides. These ratings were used to divide the patients into two morbidity groups, a high and a low. The patients were given an MCS test three times, over a five month period. The mean scores of the two levels of morbidity are shown in Table 2. Comparison in each row between the low and high morbidity shows that the score is higher (more supraordinate) for the high morbidity group of patients. The difference for the second and third testing are significant at satisfactory probability levels. The number of *Ss* given in the table are variable because of the variable cooperation of *Ss*, and in the last testing a few were not in the hospital.

SUMMARY AND CONCLUSIONS

A multiple choice method of testing supraordinate associations was used with several groups of varying educational levels, including a group of hospitalized mental patients who were the least educated. The following conclusions are supported by the results: (a) There is a greater tendency toward subordinality of associations in the multiple choice form of the test than in the free association method. (b) A positive relationship between education and supraordinality of association predominates in the multiple choice method. (c) Within groups of homogeneous educational level, there is a positive relationship between supraordinality and maladjustment.

REFERENCES

1. PETERS, H. N. Supraordinality of associations and maladjustment. *J. Psychol.*, 1952, *33*, 217-225.
2. PETERS, H. N. Supraordinality of associations and mirror-tracing ability. *J. gen. Psychol.*, 1954, *50*, 245-248.
3. PETERS, H. N. Circular pencil maze performance in chronic schizophrenics. *J. clin. Psychol.*, 1956, *12*, 170-173.

AN EXPERIMENTAL STUDY OF TRANSFERENCE AND RESISTANCE PHENOMENA IN PSYCHOANALYTICALLY ORIENTED PSYCHOTHERAPY[1]

MOSS L. RAWN

Creedmoor Institute for Psychobiological Studies
Queens Village, New York

According to Freud the two essential attributes of psychoanalysis are transference and resistance (4, p. 939), and the basic task in Freudian psychoanalysis is the continuing analysis of transference and resistance (18, pp. 90-91). Objectification and refinement of these two concepts is possible without removal from naturalistic settings. Stephenson's Q-technique (3, 5, 14, 16), already applied by Rogers and his associates (15) to test the nondirective counseling process, provides a way of examining transference and resistance phenomena in an ongoing psychoanalysis.

[1]The author wishes to express gratitude to John R. Whittier, M.D., Director of the Institute, for encouragement and suggestions.

Fenichel, following Freud, wrote "everything that prevents the patient from producing material derived from the unconcious is resistance"[8, pp. 27-29]. In the Freudian view, resistance is in essence an *active, unconcious* process that acts in *opposition* to interpretation[13, p. 246]. A resistance-present, resistance-absent bipolarity is operationally developed by setting up items that carry various attributes of each pole.

Transference is seen by Freud as an unconcious process, rapport being its stereotype[4, p. 936]. Modifying Monroe's[10, 13] formulation of the Freudian position yields transference as an *unconcious polarized emotional set* representing a repetition in the present of *past emotional states*. Within this framework a trichotomy may be conceived: positive transference or the like, love, sexual states; negative transference or the dislike, hate, hostile states; and neutral[2] transference or the calm, neutral, rational states. Items are operationally devised for transference that contain various aspects of each dimension of the trichotomy.

Hypotheses

The general hypothesis is that transference and resistance phenomena through the course of psychoanalytically oriented psychotherapy may be quantitatively identified. Furthermore, specific transference and resistance attributes for a particular patient may be quantitatively ascertained. These hypotheses are translated for testing purposes into an analysis of variance balanced block Q-design for one patient[3] resulting in the following sub-hypotheses:

1. Items may be devised which can be coded by significant inter-judge agreement into transference and resistance polarities (designing the scale).

2. There will be no significant difference among raters in Q-describing a particular patient for the transference and resistance variables (reliability of the scale).

3. The transference variance will be significant for single and for pooled sessions (identification of an attribute).

4. The resistance variance will be significant for single and for pooled sessions (identification of an attribute).

5. Transference-resistance interaction will be insignificant for single and for pooled sessions (independence of attributes).

6. There will be significant session-to-session variance differences for transference and resistance, differentiated or undifferentiated (session variation of attributes).

7. Significant transference and resistance details characteristic of a particular patient will emerge (quantitative description of an individual's psychoanalysis in terms of psychoanalytic variables).

Method

Personnel. A patient-therapist unit was chosen to typify psychoanalytically oriented psychotherapy[4] process. Five paid *raters*[5] were selected naive to the design, but sufficiently sophisticated, in the author's opinion in psychoanalytic thinking and practice to warrant their use in the experiment. They supplied the tape-listening, note-writing, and Q-sorting. Five *other* individuals, unpaid, acted as *judges* in the preparation of the Q-statements.

[2] A concept coined to include the "gray" zone between positive and negative states.

[3] Psychoanalysis examines one case at a time intensively, and happily a design is available, Q-technique, for quantitative investigation of the one case, making the method relevant to the phenomena being studied[5, 11, 12, 14, 16].

[4] Wolberg makes distinctions among the various schools of psychoanalytic and non-psychoanalytic psychotherapies[18, p. 80, 90-91] and his criteria for establishing the term "psychoanalytically oriented psychotherapy" were followed by the experimenter who selected a fitting patient-therapist combination.

[5] Four of the five raters were matriculated students at a psychoanalytic training center and the fifth was a matriculant of an analytically oriented counseling program. All had direct therapy experience with their own patients.

Development of the Scale. Allport's[1] compilation of twelve thousand emotionally toned words and other words the author selected from his own experience, provided the mass from which items were drawn in conformity to the transference and resistance polarities. Lists were made totalling several hundred items. Several of these, placed in the appropriate categories are as follows:

Transference			Resistance	
Positive	Negative	Neutral	Present	Absent
generous	belligerent	objective	falsifying	insight
cheerful	quarrelsome	composed	covert	analyzing
tender	envious	casual	evasive	uninhibited

One transference and *one* resistance item were combined to yield a transference-resistance polarized sentence. An example of such a polarized sentence is: "In an envious context, falsifying occurred." All combinations of three transference and two resistance items were composed with sufficient replication to yield a list larger than the design calls for, allowing thereby a choice of best items. The refined list contains a 3 by 2 matrix for 15 replications or 90 sentences arranged in a balanced block design illustrated in Table 1.

TABLE 1. TRANSFERENCE AND RESISTANCE CELL MATRIX

Resistance	Transference Positive	Negative	Neutral	Mean
Present	A	B	C	R
Absent	D	E	F	X
Mean	P	N	O	

Note: Cells will be designated by letters A-F inclusive; vectors by P, N, O, R, and X; and sessions by I, II, III, and IV. This matrix is used for each session.

Five judges independently coded 150 randomized sentences each into one of the 6 cells, basing their decisions on the operational definitions of transference and resistance defined above. The judges were informed that all cells were to be represented but were told neither about equal representation of cells nor about the purpose or design of the study. A sample statement for judging follows: "Although the session contained significant silences, grateful affect seemed present." This could be coded, using the definitions, into cell A (transference-positive, resistance-present).[6] Agreement of four judges out of five constituted criteria of sufficient grounds for statement retention. With six choices per statement (Cells A through F) for each of five judges, absolute agreements among four judges appeared at a P level of .02 and five out of five judges in complete agreement reaches a P of .0008.[7] More than the required 90 statements in balanced block design with significant cell placement agreement emerged out of a possible 150. The best 90 items were used with 54 statements reflecting five out of five judges in complete agreement and 36 statements exhibiting four out of five complete agreements among judges. Clearly, these statements have been successfully coded into transference-resistance polarities and may be safely used as the Q-scale defining the experimental task.

Administration. Four tape recorded sessions of one patient provided basic material. The sessions were spaced for adequate time sample over a one year period of psychoanalytically oriented psychotherapy. The period spanned the second year of treat-

[6] One might argue that conscious affect only is being considered, and not unconscious transference. This overlooks an incidental control stemming from the very nature of the experiment, viz: raters decide upon listening to a tape, how characteristic an item is for the patient. The raters' judgments are based not only upon estimation of surface affect, but upon underlying feeling as well. For example of this appreciation of unconscious cues, the raters reported in their notes that the patient's laughter is to be interpreted as hostility (not as warmth). The notes freely indicate that exhibited expression is not necessarily taken at face value.

ment, session frequency being twice weekly. The tapes were presented in random time order to five raters simultaneously in four weekly two hour meetings. The raters independently Q-sorted the prepared list of ninety transference-resistance polarity coded statements once each week for the four therapy sessions following the distribution of Table 2. Raters were told to listen only to patient content and

TABLE 2. Q-SORT DISTRIBUTION (TOTAL FREQUENCY = 90)

	"Most Like"						"Most Unlike"					
Score	11	10	9	8	7	6	5	4	3	2	1	0
Frequency	3	5	6	9	10	12	12	10	9	6	5	3

dynamics revealed in the particular session played and as much as possible, to ignore therapist contributions as well as patient material of the other sessions. No information was given to the raters about the patient or therapist, except that they were to listen to psychoanalytic sessions of a particular patient-therapist combination. Written notes taken independently during the listening period and written summaries of these notes were allowed to be used freely in the Q-sorting phase conducted immediately after the tape playing. Not until completion of the fourth or final meeting were the raters informed about any aspect of the experiment or hypotheses. All expressed completely erroneous conceptions when asked upon termination to guess purposes and hypotheses. Communication among raters about the study was nil until the study was ended.

RESULTS

Having significantly established statement sub-group transference and resistance operational definitions *before* proceeding with the experiment made possible unambiguous identification of specific transference-resistance vectors appearing as a result of the experiment. Had this pre-coding operation not been established prior to experiment, ascribing significant sub-group differences to differences in transference-resistance variables would represent an unwarranted inference.

Testing differences in variances among the five raters against replication variances for transference-resistance cell by session (table 3) yields 25 out of 30 F ratios

TABLE 3. ANALYSIS OF VARIANCE F RATIOS FOR SIGNIFICANCE OF DIFFERENCES AMONG THE FIVE RATERS**

Transference–Resistance Cell	Sessions I	II	III	IV	Overall
A	1.10	.65	1.63	.58	2.56*
B	.76	.23	2.90*	.74	.78
C	3.05*	.65	.34	1.50	1.39
D	2.89*	1.33	2.50	1.64	.88
E	.56	.33	1.27	.19	1.38
F	1.89	1.01	2.43	1.27	4.15*

*Significant at P equal to .05 or below.

**The use of the same Q-sample for each session leads to equal total variances making impossible direct comparison among such totals for differences. Comparisons may be made by sub-totals, wherein extant differences could be reflected, and the cell A through F breakdown serves this end.

not significant — a finding itself considerably better than chance would dictate (Sign test t equals 3.48, P equals .01 — .001, following a procedure outlined in Walker and Lev[17, p. 431]). It is apparent that differences among raters are negligible, the sc le achieving a measure of reliability. Accordingly, all variance testing to follow treats the five raters' data as a unit increasing the replication degrees of freedom five-fold.

TABLE 4. ANALYSIS OF VARIANCE SUMMARY FOR SESSION, TRANSFERENCE, AND RESISTANCE

Source	Sum of Squares	d.f.	Variance Estimate	F	P
A. Pooled Sessions, Transference and Resistance Undifferentiated.					
Between	2603	5	520.6	82.77	> .001*
Within	11287	1794	6.29		
Total	13890	1799	7.72		
B. Pooled Session, Transference and Resistance Differentiated.					
Between Resistance	420	1	420	66.77	> .001*
Between Transference	2156	2	1078	171.38	> .001*
Interaction - R × T	27	2	13.5	2.15	< .05
Within	11287	1794	6.29		
Total	13890	1799	7.72		
C. Transference, Resistance and Session Differentiated.					
Between Resistance	420	1	420	72.53	> .001*
Between Transference	2156	2	1078	186.18	> .001*
**Between Sessions	0	3	0		
Interaction—R × T	27	2	13.5	2.33	< .05
R × S	46	3	15.33	2.65	.06
T × S	944	6	157.33	27.17	> .001*
R × T × S	15	6	2.50	.43	< .05
Within	10282	1776	5.79		
Total	13890	1799	7.72		

*Significant at P below .001
**Comment about total session-to-session variances found in footnote for Table 3 applies here. Session-to-session differences are examined via subtotal analysis in Table 5.

Transference-Resistance-Session Hypotheses Testing. Taking an analysis of variance for all four sessions as a unit with transference and resistance variances undifferentiated (Table 4A) results in a significant "between" estimate. The pooled judgments clearly cluster into the pre-coded experimental variables. Differentiating transference and resistance from each other (Table 4B) yields both variables to be significantly greater than replication, with their interaction negligible. Transference and resistance emerge therefore, as separate variables, each quantitatively identified.

Further breakdown of the variances (Table 4C) reveals the importance of the transference-session interaction. Apparently, one must consider the specific session when thinking about transference. There are (Table 5) significant session-to-session variances: a finding one would expect *a priori*. The scale seems sensitive enough to reflect significant differences in session, transference, and resistance.

TABLE 5. SUB-TOTAL SESSION-TO-SESSION ANALYSIS OF VARIANCE F RATIO SUMMARY WITH POOLED TRANSFERENCE AND RESISTANCE

Cell	A	B	C	D	E	F	Total**
F Ratio	3.37	11.57	13.48	2.90	19.75	4.78	—
P Value	.05–.01*	> .001*	> .001*	.05*	> .001*	.01–.001*	

*Significant at P of .05 or below
**See comment on total session-to-session variances in Table 3 footnote.

Considering each session separately and utilizing transference and resistance either undifferentiated or differentiated (Table 6) further confirms the positive findings as does examination of Fig. 1. In each test the transference and resistance variables emerge not only as distinct, but also as independent.

Having established significant variance differences for the experimental variables transference, resistance, and session makes meaningful closer examination of mean cell score for characterization of an individual psychoanalysis. It can be seen

TABLE 6. ANALYSIS OF VARIANCE F RATIO SUMMARY FOR TRANSFERENCE AND RESISTANCE BY SESSION

	Session I		Session II		Session III		Session IV	
	F	P	F	P	F	P	F	P
Transference and Resistance *Un*differentiated.	33.22	> .001*	11.21	> .001*	32.16	> .001*	56.44	> .001*
Transference and Resistance Differentiated.								
Resistance	10.72	> .001*	17.40	> .001*	32.23	> .001*	7.32	.01 - .001*
Transference	77.24	> .001*	8.81	.01 - .001*	62.89	> .001*	137.34	> .001*
Interaction R × T	.44	< .05	1.51	< .05	1.39	< .05	.11	< .05

*Significant at .01 or below

(Table 7) that in the first session negative transference with or without resistance is significantly more present than all other combinations[7]. Session II mostly contains neutral transference with resistance present; session III reflects negative transference with resistance present; session IV indicates similar results as does session I;

TABLE 7. TRANSFERENCE-RESISTANCE CELL MEAN SCORE BY SESSION AND SIGNIFICANT DIFFERENCES BETWEEN MEAN SCORES

Cell	Session I	II	III	IV	Total	Significant Differences Between Mean Scores by Session*
A	4.44	5.25	4.51	4.03	4.56	I-II II-IV
B	*7.79*	6.00	*7.79*	*8.04*	*7.40*	I-II II-III II-IV
C	5.39	*7.16*	6.13	5.28	5.99	I-II I-III II-III II-IV III-IV
D	3.91	4.57	3.67	3.45	3.90	II-III II-IV
E	*7.08*	4.39	6.53	*7.59*	6.40	I-II II-III II-IV III-IV
F	4.40	5.63	4.37	4.61	4.75	I-II II-III II-IV
**Total	(5.50)	(5.50)	(5.50)	(5.50)	(5.50)	
Significant Differences Between Mean Scores by Cell*	A-B A-C A-E B-C B-D B-F C-D C-E C-F D-E E-F	A-C A-E B-C B-D B-E C-D C-E C-F D-F E-F	A-B A-C A-D A-E B-C B-D B-E B-F C-D C-F D-E E-F	A-B A-C A-E B-C B-D B-F C-D C-E D-E D-F E-F	A-B A-C A-D A-E B-C B-D B-E B-F C-D C-F D-E D-F E-F	

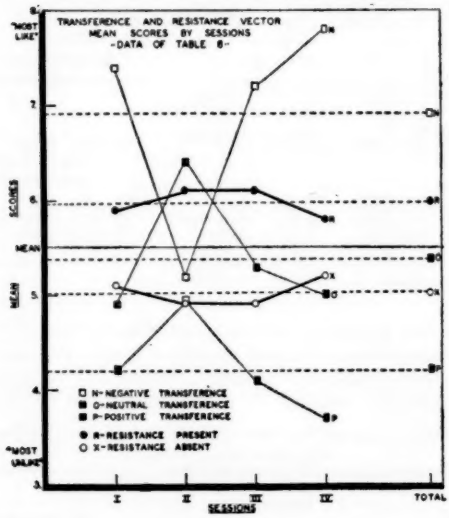
Note: Italized scores represent most prominent transference-resistance valences in a session.

*Significant at P equal to .05 or below.
**For comment about session-to-session differences among totals, see footnote on Table 3.

and for all sessions together transference is negative with resistance present. *This* patient's psychotherapy may be described in terms of the predominance of negative transference simultaneous with sizeable resistance. Furthermore, there has been a

[7]Cell B (transference negative, resistance present) mean score 7.79 is significantly larger than all other cells save cell E (transference negative, resistance absent) mean score 7.08. The latter likewise is significantly larger than all other cells except B. Cell B although larger than cell E, is not significantly so. The reader may make similar comparisons for the other sessions.

FIG. 1. TRANSFERENCE AND RESISTANCE VECTOR MEAN SCORES BY SESSIONS (From Table 8)



quantitatively determined ebb and flow of transference and resistance effects (Fig. 1 and Table 8) throughout the four sessions sampling the analysis. Negative transference gives way to a more neutral state only to re-emerge, becoming at the end even more negative than ever before. The transference curves show distinct changes from session to session. The resistance curves, however, are flatter, revealing steady maintenance of considerable resistance.

TABLE 8. TRANSFERENCE AND RESISTANCE VECTOR MEAN SCORES BY SESSION

Vector	Session I	II	III	IV	Total
Transference					
Positive P	4.17	4.91	4.09	3.74	4.23
Negative N	7.43	5.19	7.16	7.81	6.90
Neutral O	4.89	6.39	5.25	4.95	5.37
Resistance					
Present R	5.87	6.14	6.14	5.78	5.98
Absent X	5.13	4.86	4.86	5.22	5.02

Note: Cells were combined to yield vectors.

CONCLUSIONS AND IMPLICATIONS

1. A reliable and consistent exploratory scale was devised that shows promise for measuring ongoing psychoanalytically oriented psychotherapy in a naturalistic setting without interfering with the therapeutic experience.[8] Furthermore, terms in keeping with psychoanalytic convention were used without calling for strained and artificial expressions foreign to psychoanalysis.

2. The terms, transference and resistance, characteristic of psychoanalytic psychotherapy, were employed and given quantitative description.

[8]Since neither patient nor therapist was aware of research design at the time of recording, the only material disturbance of the therapeutic process was the presence of a microphone. The author in agreement with Rogers[15] does not feel this was a substantial interference.

3. Transference and resistance were found to be separate and distinct variables. This independence of transference and resistance suggests the presence of intervening variables, *viz.*, patient's character structure—possibilities rich for future study.

4. Samples of a patient's psychotherapy were given quantitative description in terms of transference and resistance details. Further description of these details in qualitative as well as quantitative ways will appear in a subsequent publication.

5. Transference and resistance phenomena were shown to vary from session to session, as might be expected. It becomes possible in future investigations to attempt to correlate these changes with process as well as outcome indices.

6. The consistently negative transference shown by the patient (except in Session II) combined with continued resistance might signify a failing therapeutic process, and examination of the patient's history shows increasing maladaptation, with accumulation of symptoms. This concordance incidental to the present study may indicate the value of the method for further exploration.

References

1. Allport, G. and Odbert, H. Trait-names, a psycho-lexical study. *Psychol. Monogr.*, 1936, *47*, #211.
2. Auld, F. and Murray, E. J. Content analysis studies of psychotherapy. *Psychol. Bull.*, 1955, *52*, pp. 377-395.
3. Bordin, E. S. Factor Analysis in experimental design in clinical and social psychology. *Psychol. Rev.*, 1943, *50*, 415-429.
4. Brill, A. A. *The basic writings of Sigmund Freud.* New York: Random House, 1938.
5. Cronbach, L. Correlations between persons as a research tool. In Mowrer, O. H. *Psychotherapy, theory and research.* New York: Ronald Press, 1953.
6. Dollard, J. and Miller, N. *Personality and psychotherapy.* New York: McGraw Hill, 1950.
7. Feller, W. *An Introduction to probability theory and its applications. Volume one.* New York: Wiley, 1950.
8. Fenichel, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.
9. Fisher, R. A. *The design of experiments.* Edinburgh: Oliver & Boyd, 1935.
10. Freud, S. The dynamics of the transference. *Collected Papers, Vol. II.* London: Hogarth Press, 1924.
11. Kantor, J. R. *Psychology and logic.* Vol. I & II. Bloomington Indiana: Principia Press, 1951.
12. Kaufmann, F. *Methodology of the social sciences.* New York: Oxford Univ. Press, 1944.
13. Monroe, R. I. *Schools of psychoanalytic thought.* New York: Dryden Press, 1955.
14. Mowrer, O. H. Q-technique description, history and critique. *Psychotherapy, theory and research.* New York: Ronald Press, 1953.
15. Rogers, C. and Dymond, R. *Psychotherapy and personality change.* Chicago: University of Chicago Press, 1954.
16. Stephenson, W. *The study of behavior, Q-technique and its methodology.* Chicago: University of Chicago Press, 1953.
17. Walker, H. and Lev, J. *Statistical inference.* New York: Holt, 1953.
18. Wolberg, L. *The technique of psychotherapy.* New York: Grune and Stratton, 1954.

INSIGHT AND PROGNOSIS

ALVIS ESKEY

Cleveland Psychiatric Institute and Hospital

It has been reported that the prognosis is more favorable for psychotic patients who show insight into the fact that their thinking is or may be distorted than it is for patients who lack such insight.[1] Henderson and Gillespie[4] state that ". . . a person who entertains a false belief, but at the same time admits it may be unjustified, and who may spontaneously talk of it as a delusion, is as a rule, more amenable to treatment than one whose beliefs admit of no such criticism." Many, if not all, of the initial psychiatric interviews given patients immediately upon their admission to a hospital, attempt to establish the presence or absence of insight. On an anecdotal basis the prognosis is often considered better for the patients who reveal evidence of insight into the fact that they are experiencing a psychological disturbance. In the present investigation an attempt was made to discover whether psychotic patients who had insight into their illness had a better prognosis (in terms of length of hospitalization) than did the patients who possessed only partial or no insight. Operationally defined, insight refers to the presence or absence of verbalized *awareness* on the part of the patient that impairment of intellectual functioning existed.

Method

Three hundred psychotic patients who at the time of hospitalization displayed varying levels of insight or who, conversely, were completely lacking in insight were selected for the study. The presence or absence of insight was determined through an evaluation of each patient's mental status examination. The age range was 15 through 70 and all patients had been hospitalized at least two weeks. Persons in the following categories were not used in the study: Those with a history of previous mental illness, organic, alcoholic, and mentally defective patients, chronic patients transferred to other institutions, individuals unimproved at the time of discharge as well as those leaving against medical advice, and patients who for any reason were unable to freely communicate their thoughts. Factors related to the composition of the groups are summarized in Table 1.

In order to compare differences in the length of hospitalization, the patients were divided into three groups, each having an N of 100. Those patients who displayed evidence of marked insight were assigned to the first group. At the time of their admission to the hospital, these individuals verbalized statements which were felt to reflect awareness of malfunctioning on their part. Such declarations as "seem confused," "something is wrong with my mind," etc., are representative of ideas expressed by members of this group. The second group was composed of patients who appeared to have at least partial insight into the fact that they were experiencing some sort of psychological difficulty. They were ambivalent in their attitudes as exemplified by the statement "don't think I need any psychiatric help, but if I do, I'll do anything to get well." Thus, while expressing indecision about the adequacy of their functioning, they were not entirely convinced that they were actually ill. Many of them expressed the opinion that "something" might be wrong with their minds, but that they were not certain of it. Patients who exhibited no evidence of insight into their illness were placed in the third group. They consistently denied the possibility of mental or emotional difficulties and in general felt that they had been hospitalized without adequate reason.

The following comparisons, in terms of the number of days of hospitalization, were made between the groups: Insight versus no insight; insight versus partial insight; and partial insight versus no insight. In addition, it was felt that the *type* of admission to the hospital might indirectly reflect the presence or absence of some degree of insight. It was reasoned that, in general, patients who came to the hospital on a voluntary basis were likely to possess more ego-intactness than those who were

TABLE 1. FACTORS SHOWING THE COMPOSITION OF THE EXPERIMENTAL GROUPS

	Insight	Partial Insight	No Insight
Sex			
Male	41	38	37
Female	59	62	63
Total	100	100	100
Marital Status			
Married	47	56	53
Single	36	25	26
Separated	3	5	10
Divorced	7	10	8
Widow	7	3	3
Widower	0	1	0
Total	100	100	100
Diagnosis			
Schizophrenia			
Paranoid	31	36	42
Catatonic	4	5	19
Simple	4	5	2
Hebephrenic	0	0	2
Chronic Undiff.	9	6	8
Acute Undiff.	7	5	3
Schizo-Affective	7	8	2
Manic	0	0	2
Depressed	14	13	3
Involutional	24	22	17
Total	100	100	100
Treatment			
ECT	72	80	85
Insulin	6	4	0
ECT and Insulin	5	3	4
Other	17	13	11
Total	100	100	100

brought in as emergency cases or who were probated by the court. In order to test this hypothesis, the following groups were investigated: Voluntary patients (N = 75) versus emergency admissions (N = 75); voluntary patients (N = 40) versus court placements (N = 40); and voluntary patients showing insight (N = 45) versus court placements with no insight (N = 45).

The length of hospitalization was used as the criterion of improvement, since it was the most readily available objective measure of patient recovery. Prognosis, in the sense in which it is used in this study, refers to how quickly the patients progressed to the point where discharge from the hospital as "recovered" or "improved" was made. The *t* test employing medians was used to determine the significance of the differences.

RESULTS

The significance of the differences between the insight groups is summarized in Table 2. On the basis of this sample group, no reliable differences in the rate of recovery were found to exist between patients with insight and those with no insight. This finding suggests that patients with insight do not improve more rapidly (as evaluated by the number of days of hospitalization) than do patients with insight. When the groups displaying the greatest extremes of symptomatology (*i.e.*, types of hospital admissions combined with the presence or absence of insight) were compared, no significant differences were discovered. The statistical results obtained from comparison of these groups are presented in Table 2.

TABLE 2. SIGNIFICANCE OF DIFFERENCES IN LENGTH OF HOSPITALIZATION BETWEEN GROUPS STUDIED FOR INSIGHT AND TYPE OF HOSPITAL ADMISSION

Group 1	Group 2	Median (Days) Group 1	Median (Days) Group 2	t	p
Insight	No Insight	75.37	80.30	.42	.67
Insight	Partial Insight	75.37	82.62	.65	.52
Partial Insight	No Insight	82.62	80.30	.20	.84
Voluntary	Emergency	73.63	74.50	.07	.95
Voluntary	Court Placement	72.47	86.39	1.02	.31
Voluntary (Insight)	Court Placement (No Insight)	70.15	87.84	.92	.36

DISCUSSION

Mere knowledge by the patient that deviant psychological functioning is present does not in itself seem to be prognostically significant insofar as length of hospitalization is concerned. This appears to be similar to the often seen psychotherapeutic situation in which progress on the part of the patient does not necessarily follow intellectual insight into the existence of the problem. It should be emphasized, however, that the kind of insight considered in this study is not the counterpart of psychotherapeutic insight wherein the patient may experience an emotional as well as an intellectual reorientation in the course of his treatment. While therapeutic insight generally is dependent upon the development of conscious awareness and understanding of formerly unrecognized impulses and drives, insight as considered in this study reflects only a verbalized awareness by the patient that some sort of psychological deficit existed.

It is possible that the lack of differences between the groups which were studied may in part be due to a superficiality of the insight which was displayed or perhaps even to suppression or deliberate distortion on the part of the patient. In some cases external sources (*e.g.*, relatives or friends) may have constituted the prime impetus which led the patient to seek hospitalization. Again, the admission that psychological difficulties existed might have in some instances been an incorporation, by the patient, of opinions expressed by other persons. Some individuals who were cognizant of their position as "mental" patients may have attempted to present a facade of normality (*i.e.*, denied being ill) while at the same time being covertly aware of their malfunctioning.

If insight is considered as one aspect of a larger complex of factors which reflect ego-intactness, then the present investigation is in agreement with previous studies which report intellectual disorganization and general confusion as having little or no effect on the speed of recovery among psychotic patients.[2, 3] It may be that the isolated personality variables which were investigated in this and other studies are not sufficiently representative of the entire personality structure to constitute valid predictors of prognosis. Further studies which utilize larger groups of factors felt to be representative of ego-intactness may prove valuable.

SUMMARY

The present study attempts to discover whether there is a significant relationship between insight (*i.e.*, verbal recognition by the patient of existing psychological difficulties) and prognosis (as evaluated by length of hospitalization) insofar as psychotic patients are concerned. The experimental population consisted of 300 psychotic patients divided equally into the following groups: Patients showing insight into the fact that they were experiencing psychological troubles; patients who had only partial insight; and patients who had no insight into their illness. No significant relationships relevant to prognosis, as evaluated by the number of days spent in the hospital, were found between any of the groups. When other objective criteria were taken into consideration along with insight (*i.e.*, the type of hospital admission) there again were no significant relationships noted. It was suggested that

these results bear a similarity to psychotherapy in which progress does not necessarily follow intellectual recognition of existing problems. The kind of insight with which this study is concerned was differentiated from psychotherapeutic insight. Factors which might have contributed to the lack of differences between the insight and non-insight groups were discussed. It was suggested that limited aspects of the personality when studied in isolation may not encompass a sufficient segment of the entire personality structure to serve as valid indicators of prognosis. Further studies were suggested in which larger clusters of factors reflecting ego-intactness might be incorporated.

References

1. Chase, L. S. and Silverman, S. Prognostic criteria in schizophrenia. *Amer. J. Psychiat.*, 1941-42, *98*, 360-368.
2. Eskey, A., Friedman, Gladys, and Friedman, I. Disorientation as a prognostic criterion. *J. consult. Psychol.*, 1957, *21*, 149-151.
3. Eskey, A., and Friedman, I. The prognostic significance of certain behavioral variables. *J. consult. Psychol.*, 1958, 22.
4. Henderson, D., and Gillespie, R. *A Textbook of Psychiatry.* London: Oxford University Press, 1952.

A COMPARATIVE JUDGMENTAL STUDY OF "DEFECTION" FROM TREATMENT AT A CHILD GUIDANCE CLINIC[1]

EUGENE E. LEVITT

Indiana University Medical Center

Problem

One of the most persistent and troublesome problems of the child guidance clinic is the family which voluntarily breaks off contact with the clinic after its staff has invested time and effort in the case. The most wasteful type of premature termination occurs at a point following a complete diagnostic, a conference staffing, and the assignment of a therapist. After the clinic has thus spent as much as twenty-five hours on the case, the patient refuses treatment. Some clinics have reported that "defectors" from treatment make up more than 30% of all cases accepted for therapy in a given period.

At present, the factors making for defection are largely unknown. It is conventional to subsume them under the heading of "insufficient motivation for treatment," and to assume that this descriptive label is explanatory.

There are few experimental investigations bearing on the question of therapy defection at the child guidance clinic. Two follow-up studies[1, 3] show that the parents of defector cases offer a wide variety of reasons for defection, suggesting that there is no unitary explanation of the phenomenon. They also indicate that improvement in the child's symptoms is a factor in from 15 to 35% of the cases. A recent comparative study[2] of defector and treatment cases among child clinic patients disclosed no differences on sixty-one diagnostic and descriptive variables.

The present investigation studies three questions pertaining to defection from treatment at the child guidance clinic: Do defectors differ from treatment cases with respect to "motivation for treatment?" Do they differ with respect to severity of symptoms (also frequently regarded by clinicians as a predictive index)? Can expert judges distinguish between defectors and therapy cases using a molar estimate based on diagnostic and identifying data?

[1]This study was supported by funds made available by the Institute for Juvenile Research of Chicago, an agency of the Illinois Department of Public Welfare.

Method

For our purposes, a defector case is defined as one which was accepted for treatment after a complete diagnostic work-up, and which failed to begin therapy. A treatment case is one in which at least one family member had at least five treatment interviews.

The records of fifty defector and fifty treatment cases were selected at random from among those cases closed at the Institute for Juvenile Research in Chicago during the period 1951-1953. All notations concerning treatment were removed from the records of the treated cases.

The records were then submitted to three expert judges[2], two psychologists and a psychiatric social worker. All were ex-IJR employees and all had had at least three years of professional experience. The judges were first asked to identify records of cases with which they had had some contact, or in which some crucial hint concerning disposition had been inadvertently left. Four defector cases and seven treatment cases were discarded on these bases. Of the remaining 43 treatment cases, twenty-four had had a multiple focus of therapy. In sixteen cases, only the mother was treated, and in three cases, only the child. The mean number of treatment interviews for the mothers was 22.29, and for the children, 24.35. The mean number of interviews per family was 36.02. In twenty-seven cases, at least one family member had had at least twenty treatment interviews, and in only six cases was there no family member with at least ten.

The judges were told that there were both defector and treatment cases in the group of records, but they were not told the relative proportions of each. The judges were then asked to make the following judgments: (a) Rate the parents' motivation for treatment for the child on a five-point scale, on which 1 was least motivated and 5, most motivated. (b) Rate the severity of the presenting symptoms in the child on a similar scale[3]. (c) Sort the 89 records into those of cases which had come in for treatment, and those which were defectors.

Results

Overall reliability of the judgments for the 89 cases was estimated by Hoyt's analysis of variance method[4]. The coefficients were .75 for motivation, .79 for severity, and .69 for the sort into defector and treatment cases. (There was 100% agreement on the sort for 52 of the 89 records.) While these reliabilities are not high, they appear adequate for the study.

The judgments for motivation and severity of symptoms were then subjected to variance analyses to determine primarily if there were differences between the treated and defector groups. A variance due to judges was also computed. The analyses are summarized in Table 1. The mean scores (averages for the three judges) for motivation are 2.98 for the treated group, and 2.88 for the defectors. The F-ratio is 0.47, which is insignificant. The F for differences among judges is 0.19, which is also insignificant. The interaction between groups and judges yields an insignificant F of 0.14. The mean scores for severity of symptoms are 2.98 for the treated group and 3.02 for the defectors. The F-ratios of 0.20 for groups and 0.14 for interaction are insignificant. The F of the differences among judges is 7.51, which has a p-value beyond .001.

---

[2]The author is greatly indebted to Drs. Libby Siegel and Rita Weinberg, and to Mrs. Marion Johnstone, who served as the judges for this study.

[3]These scales were admittedly *ad hoc*, and no great pains were taken to maximize their interjudge reliabilities. It was assumed that the commonality of experience at IJR would itself have served as a training period for the judges for the tasks required by the present study. However, an orientation session was held, and certain cues were provided as bases of the judgments. It was suggested that indicators of low motivation might be the parent's statement that coming to the clinic was not her own idea, expressions of doubt about the efficacy of treatment, and reluctance on the part of the parent to become a treatment focus; there might be a record of broken appointments, and of previous clinic treatment, either at IJR or elsewhere. Four major factors were to be considered in evaluating severity of the disturbance: its scope in the child's life, its chronicity, atypicality (i. e., degree of bizarreness), and intensity (degree of incapacitation).

TABLE 1. ANALYSIS OF VARIANCE OF JUDGMENTS OF MOTIVATION AND SEVERITY OF SYMPTOMS OF DEFECTOR AND TREATMENT SAMPLES

Source	Motivation			Severity of Symptoms		
	df	Mean Square	*F*	*df*	Mean Square	*F*
Judges	2	0.260	0.19	2	5.270	7.51*
Samples	1	0.660	0.47	1	0.140	0.20
J x S	2	0.200	0.14	2	0.095	0.14
Within cells	261	1.392		261	0.702	

*P = .001

A summary of the results of the judges' sort into treatment and defector cases is shown in Table 2. All judges identified treated cases better than chance at the .01

TABLE 2. PERCENTAGES OF CASES IDENTIFIED CORRECTLY BY THE JUDGES

Groups	Judge A	Judge B	Judge C	Average
Treated Cases (N = 43)	72.1**	81.4**	88.4**	80.6**
Defectors (N = 46)	58.7	56.5	37.0	50.7
All Cases (N = 89)	65.2**	68.5**	61.8*	65.2**

*Significantly different from chance at the .05 level
**Significantly different from chance at the .01 level

level. The average percentage of correctly identified treated cases is 80.6. However, no judge identified defector cases better than chance, the average being only 50.7%. The overall percentage of cases correctly identified by the judges, the most adequate index of accuracy, is 65.2. This is significantly better than chance at the .01 level. The range of over-all percentages per judge, which is about 61% to 69%, suggests that the overall average is a fairly stable estimate of accuracy.

The probable influence of the judges' estimates of motivation and severity on the sort into treated and defector cases is a matter of incidental interest. These influences can be investigated by means of biserial correlations between each judge's rating of motivation or severity and her own sort. The biserial coefficients are shown in Table 3. All of the correlations between motivation ratings and sorts are

TABLE 3. BISERIAL CORRELATIONS BETWEEN THE JUDGMENTAL VARIABLES AND THE SORT INTO DEFECTOR AND TREATED CASES FOR EACH JUDGE

Between the Sort and	Judge A	Judge B	Judge C	Average
Motivation for treatment	.66**	.91**	.90**	.85**
Severity of symptoms	.00	.31*	.22	.18

*Significant beyond the .05 level
**Significant beyond the .001 level

significant beyond the .001 level. Each judge tended to ascribe higher motivation to the cases which she judged as having entered treatment. However, the correlation between severity judgments and sorts was significant (P = .05) only for Judge B, who tended to rate the defector cases as having greater severity.

DISCUSSION

The data of Table 1 demonstrate that the defector and treated samples did not differ along the dimensions of severity of symptoms or motivation for treatment as

these variables were defined in this study. Although the judges differed significantly among themselves in mean ratings on the former variable, the absence of interaction between judges and samples (even the absolute magnitude is quite small) indicates that these differences are systematic over both samples for all judges. Mean ratings of the judges on motivation did not differ and the interaction is again small and insignificant.

The fairly high accuracy of identification of treated cases by the judges is in part a function of the fact that all judges judged that the percentage of treated cases of all cases was higher than it actually was. This is reflected in the inverse relationship between accuracy of identification of treated and of defector cases. The greater the judge's accuracy for treated cases, the lower her accuracy for defectors. Notably, no judge identified defector cases better than chance. The overall percentages of correct identification of cases (65.2, 68.5 and 61.8) are all better than chance, but they are not very impressive nonetheless. Furthermore, the important action from the point of view of clinic operation is not to identify the good treatment risks, but to single out the prospective defectors. Viewed from that angle, the judges were entirely unsuccessful

The data of Table 3 show, not surprisingly, that the judges based their sorts largely upon their judgment of motivation for treatment. The magnitude of the correlations suggests that they hardly considered any factor other than motivation in deciding if the case would come for treatment. This clearly reflects the contemporary theoretical view of the significance of motivation in determining defection or continuance in therapy, despite the lack of clarity of the concept, and the paucity of experimental evidence. Severity of symptoms appears to have influenced the sorts minimally, or not at all.

Conclusions

The findings of this study suggest that motivation for treatment and severity of symptoms as judged by expert clinicians on the basis of diagnostic and identifying information do not differentiate treatment from defector cases at the child guidance clinic. There is also experimental confirmation of the fact that the clinic staff appears to be relatively unsuccessful in identifying prospective defector cases on the basis of diagnostic and identifying data.

An alternative inference would be that expert clinicians are presently unable to clinically rate motivation for treatment and or severity of symptoms with adequate accuracy. By accepting such an interpretation, one may retain the hypothesis that either or both of these factors actually distinguishes the defector case from the treatment case. The data of this study do not clearly support either of the alternative interpretations.

References

1. Inman, Ann C. Attrition in child guidance: a telephone follow-up study. *Smith Coll. st. soc. Wk.*, 1956, *27*, 34-73.
2. Levitt, E. E. A comparison of "remainers" and "defectors" among child clinic patients. *J. consult. Psychol.*, 1957, *21*, 316.
3. Levitt, E. E. Parents' explanations of "defection" from treatment at a child guidance clinic. *Ment. Hyg.*, in press.
4. Lindquist, E. L. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

ALTERNATING TWO-DAY CYCLIC BEHAVIOR CHANGES

A. A. WEISS*

Psychiatric Hospital "Talbieh", Jerusalem

INTRODUCTION

Two-day cycles of alternating "good" and "bad" behavior in psychiatric patients are a rare phenomenon. Menninger-Lerchenthal, reviewing the literature in 1950 [4], cited 52 cases reported since 1798 by a total of 34 authors. Richter [5] reviews the behavioral features of a group of such patients observed by him. While several of his patients are in an objectively better state on their "good" days than on "bad" days, this is by no means true of all his patients. Such patients may be gay, cheerful, contented, confident, elated on their "good" days and depressed, discouraged, hypochondriacal and fearful, irritable or listless or apathetic on their "bad" days. They tend to be talkative, enterprising, full of ideas, active and not easily tired as well as cooperative on their "good" days, but some of them are uncooperative and even negativistic, rigid, weak and inactive, sleepy and repetitious, and even stuporous on their "bad" days. Others again may be restless, distractible, agitated and even destructive on "bad" days. They may be normal and rational, but also incoherent and boastful on "good" days, while showing poor judgment and a tendency to worry on "bad" days. Some of them may actively hallucinate on "good" days, while showing marked restlessness on "bad" days. Some patients have gastric pains, increased pulse and electrical skin resistance, high temperature, increased salivation, dull eyes and dull immobile facial expression on "bad" days. Thus, behavioral features and mood do not follow a unitary pattern, one patient appearing more disturbed on "bad" days while more or less "normal" on "good" days, whereas another patient reverts to a tolerable level on his "bad" day after having shown hypomanic features on his "good" day.

The phenomenon was initially considered to be a special form of a cyclic affection of the manic-depressive type. However, as early as 1929, Arndt [1] stated that "the daily change of phase in itself has nothing to do with the manic-depressive circular affection." More recent investigations and attempts at explanation concern —apart from the observation of behavior and mood swings—the neurological, physiological, metabolic [6] aspects of the illness, while no psychodynamic or psychodiagnostic study could be found in the extensive literature quoted by either Menninger-Lerchenthal or Richter, or in journals available to the present writer. It should, therefore, be of special interest to give an account of the observations made on a male patient aged 42, married and father of 4 children who was hospitalized at his own request on the open ward of "Talbieh" Psychiatric Hospital for 30 days. Our account will comprise the clinical observations, physical, psychiatric and psychodiagnostic examinations and an outline of his life history.

Life History. Pt. was born in Czarist Russia as the second son of a well-to-do Jewish timber merchant. When he was 4, his father died from cancer at age 37. After World War I the mother and her 2 sons rejoined her family in Poland. His mother did not remarry until pt. was 24. At the time of his hospitalization she was reported to be healthy at the age of 72. Pt. finished high school at 19. He was a good scholar who earned some money by coaching while still at school. He was an active Zionist and emigrated with his mother and brother to Palestine at the age of 20. He joined a collective settlement where he quickly rose to leading positions. At 28 he married. His public duties entailed a good deal of travel while his wife and children remained at home. At the time of hospitalization he worked as a leader of pioneering groups spending varying periods away from home. Pt. suffered from ricketts and pertussis as a child; at 15 he is reported to have had TBC.

---

*Chief Clinical Psychologist, Psychiatric Hospital "Talbieh", affiliated to the Hebrew University Medical School, Jerusalem; Clinical Psychologist, Dept. of Neurol., Hadassah Hebrew University Hospital, Jerusalem.

Onset of Present Illness. When pt. was 36, his youngest son, then 2 mos. old, was hospitalized. While away from home, pt. erroneously was notified that his son was dying. In fright he rushed to the hospital to discover the error. He donated blood, lost 18 kgs. of weight in 10 weeks and developed a tremor in his hands. In 1950 he was hospitalized and found to have a BMR of +72. After treatment with Propyl-Thiuracyl he felt well, working hard until 1953 when his present condition started. The evening before setting out on a trip with the youth group of which he was in charge, he participated very actively in a discussion concerning the expulsion of a member of his settlement. He then felt very weak both physically and mentally. He confirmed that he was emotionally deeply involved in the pending decision. The day after he felt well, set out on the trip. But the following day he felt poorly, pessimistic, wanted to lie down. As he felt well again on the third day, he thought that his weakness was due to lack of sleep. Only much later did he notice the regular sequence of "good" and "bad" days. ECG and BMR were normal at that time. Radioactive iodine uptake was negative. No evidence of circular mood changes in the past nor any heredity could be found. Pt. reported, however, that aggressive tendencies had long posed problems. EEG was without any pathological features.

Physical Examination. Pt. is of asthenic-athletic build with a slightly high chest. Thyroid was possibly slightly enlarged. A slight tendency to hernia on the left, and a slight contraction of the fingers, and the little finger in particular, more so on the left hand, was found. This is familial, his brother also showing such contractions. He is reported to be suffering from similar mental symptoms and a BMR of +50. However, this brother was not seen in our institution.

Psychiatric examination. Hearing, vision, and concentration were poor on "bad" days. Pt. had no appetite, slept a lot, felt as if he had saliva in his throat. He could not read serious stuff, had no sexual desire, but suffered from suicidal ideas. On "good" days he saw and heard well, was in a good mood, ate a lot, wanted to work, could not sleep. He tended to get excited over trifles. He said that on "good" days he was "his natural self plus 30%". He was hypomanic, talked incessantly, could not stick to one subject. He was critical of his physician, lacked inhibitions up to loss of decorum. He complained that although he was able to arrange his work in such a manner that the "bad" days, which were predictable in accordance with the calendar, did not interfere, he did not see how he could continue for long doing the work of two days in one. He was diagnosed clinically as a case of "short cycloid waves of the manic-depressive type superimposed upon a neurotic character of the narcissistic type".

Rorschach Results

A Rorschach test was administered on a "bad" day and repeated 5 days later on a "good" day. The scoring follows the Swiss usage with special reference to Bohm's textbook[2]. The two records will be compared under the headings of productivity, utilization of intellectual endowment, ego strength in terms of ability to control and canalize impulses and affects, affect, social adaptation, aggression, intellectual viscosity and stereotypy, insecurity and feelings of insufficiency, dysphoric mood swings.

Productivity. The number of responses (R) – exceeding normal range even on the "bad" day – increases on the "good" day by no less than 64%. Average reaction time (T/R) is short on either day. In pt's approach to the experimental stimuli attention to small details (Dd) and space details (Dzw) increases considerably, reflecting both pedantic circumstantiality and a strong oppositional tendency. Ability for the inner working through of experiences (B) appears to be non-existent in either record, while the need for the expression of (unstable and egocentric) affect (FbF) grows. The tendency towards dysphoric mood swings (HdF) is more marked on the "good" days.

TABLE 1. COMPARISON OF RORSCHACH RESPONSES ON "BAD" AND "GOOD" DAYS. ADDITIONALS ARE MARKED [ ]; G + 2(7) MEANS 2 GOOD G OUT OF A TOTAL OF 7 G.

Scoring Categories	"Bad" Day	Good "Day"
	PRODUCTIVITY	
R	23 [1] + 16 = 39 [1]	40 [4] + 24 [1] = 64 [5]
T/R	26″ : 23″ : 25″	24″ : 18″ : 22″
Erfass. Typ. (Approach)	G-Gzw-DG-*D*-Dd-DdD-*DdzwD*	G-*D*-Do-*Dzw*-*Dd*-*Ddzw*-DdD-DdzwD
Prim. Erleb. Typ. (Exp. Bal.)	0 : 6 [1½]	0 : 7 [½]
Sek. Erleb. Typ " "	0 : 1½	0 : 2½
	UTILIZATION OF ORIGINAL ENDOWMENT	
F+%	73	66
Sukz. (sequence)	gelockert (loose)	gelockert (loose)
G+ (W)	2(7)	1(7)
Erfass. Typ. (Approach)	as above	as above
T% (A%)	33%	28%
Sex%	0	18%
Obj%	20% } 53%	18% } 75%
Geograph%	0	11%
Vulgar (P)	I, II, III, V, VIII 2x	II, [III], V, VIII, IX
Orig+	0	1
Orig−	1[1]	2
B+ (M)	0	0
	EGO STRENGTH = CONTROLS	
F+%	73	66
G+ (W)	2(7)	1(7)
B+ (M)	0	0
	AFFECT	
FFb+ (FC)	2	0 [1]
FbF (CF)	5	7
Fb (C)	0 [1½]	0
	SOCIAL ADAPTATION	
Vulg (P)	6	4 [1]
D+	19 (24)	26 (41)
FFb+ (FC)	2	0 [1]
M:Md (H:Hd)	1:4	1:3
Erleb. Typ. (Exp. Bal.)	0:6 [1½]	0:7 [½]
	AGGRESSION	
Gzw− (Ws)	1	0
Dzw+ (Ds)	0	0
Dzw	0	2
Ddzw+ (Dds)	0	2
Ddzw−	0	1
DdzwD+	1	0
DdzwD	1	0
DdzwD−	1	1
Deskription	0	6
	IDEATIONAL VISCOSITY & STEREOTYPY	
Grobe Perseveration	3	20
Kleben a/Thema	4	2
Perseverat. d. erfasst. Teile	6	7
Percept. Perseveration	1	0
T%+Sex%+Obj.%+Geogr.%	53	75
	INSECURITY	
S. Kritik	1 [1]	0
O. Kritik	4 [6]	5
Symmetrie	2	8
Oder Antwort (Alternative)	1	1
Perspektive	2	1
Mittenbetonung (Midline)	0	10
Verneinung (Denial)	0 [1]	0
Geographie	1	7
Pflanzen (Bot.)	0	1
	DYSPHORIC TENDENCY	
FHd+ (Shading)	1	0
HdF	1	2

Utilization of Original Endowment. Pt's present level of intellectual functioning is not really satisfactory. His level of sharp perception (F+%) is no more than average on his "bad" days and falls off on his "good" day. In spite of a tendency to grasp complex situations, their intellectual evaluation is often unsatisfactory (few good G) in particular in consequence of emotional factors. Pt. is circumstantial (Dd), but at the same time reveals a tendency to draw rash conclusions from small details (DG, DdD, DdzwG). His thinking is stereotyped under the impact of complexive factors (Sex%, Obj%, Geogr.%) rather than showing true stability (T% low). While marked anal traits (Obj%) are found in either record, pt. tends to "show off" (Geogr.%) on his "good" days. The appearance of "sex" responses on his "good" day would indicate a marked relaxation of the critical evaluation of his own behavior. Adaptation to commonly accepted trends of thought (V) is satisfactory on his "bad" day, but less so on his "good" day. His thinking lacks depth (B), originality, and sequence (Sukzession). In view of pt's scholastic achievements and his social position utilization of his original intellectual endowment cannot be considered satisfactory.

Ego Strength as measured by the level of sharp perception (F+%), the limited ability to grasp complex situations correctly (few good G) and to work through experiences (B) must be assessed to be unsatisfactory on the "bad" day and poorer still on the "good" day.

Affects would appear largely unstable and egocentric (FbF) on either day. However, two responses indicating some measure of emotional adjustment (FFb) disappear on the "good" day, their place being taken by two more responses of the infantile, unstable type (FbF).

Social Adaptation. Interest in human beings is very limited (among a total of 103 responses obtained in both records there are no more than 2 "Human" M). Pt. appears to be fearful of contact with people (M < Md). Ability for emotional contact (FFb) disappears on the "good" day, reality testing and adaptation to common trends of thought is poorer. Pressure of aggressive trends increases. 4 responses containing the element of space (zw) increase to 7 such responses on the "good" day. Repressed aggressive energy as represented by "descriptions" shoots up from nil to 6. Simultaneously, insecurity increases markedly on the "good" day. Pt. looks for external support (emphasis on "symmetry" and "midline"). An unusual reinforcement of perseverative tendencies in his thinking would bear out the general conclusion that on what he conceives to be his "good" day, pt. is really in much poorer condition. There is evidence supporting the contention that pt. belongs to that group of people distinguished by marked intellectual viscosity: Responses reflecting the tendency to stick to certain categories such as "heads of animals", perseveration of the blot section interpreted, perceptional perseveration[2]. However, over and above these indicators, gross perseveration increases sevenfold on the "good" day comprising almost one-third of total production. From the contents of his responses we learn that in the eyes of our pt. sexual demands are closely allied with aggressive tendencies ("sex" responses being closely followed by "swords", "bombs" etc.).

Discussion

Diagnosis. Upon reviewing the two Rorschach records it becomes clear that while pt's test behavior on his "good" day might be interpreted as hypomanic, there is nothing to justify the assumption of a true depression on his "bad" day in spite of certain features in his clinical behavior. Productivity, poor intellectual control, unstable affects and a number of other features would preclude an interpretation in the depressive direction. Rather are we dealing with a neurotic and markedly narcissistic personality showing a good deal of intellectual viscosity. Ego strength is unsatisfactory and so is social adaptation.

Symptomatology. A most surprising feature in our case was the discrepancy between pt's subjective evaluation of alternate days and the psychodiagnostic findings which, however, were corroborated by the impressions gained from observation on the ward. Pt's evaluation appears to refer to his ability to work and enjoy himself with little regard to, or ability for critical evaluation of, the impact of his behavior on his environment.

Dynamic Relationship. Comparison of the two records revealed a weakening of ego control on the "good" day coupled with the emergence of more massive indications of sexual phantasies allied with considerable aggressive tendencies. This shift on the psycho-diagnostic level is correlated with a shift in certain ego functions such as vision, hearing, concentration, work, sexual activity, etc., on the clinical level. This shift raises the question, apart from the basic problem of the origin of the shift, as to whether we are dealing with a primary weakening of the Ego and a *relative* release of energy; or, conversely, whether we are confronted with a *relative* weakening of the Ego in consequence of increased instinctual pressure.

Summary

This paper gives a short review of the literature on the syndrome of alternating two-day cyclic behavior changes. An illustrative case presents the life history, clinical findings and test results of a 42 year old male with a history of two day cyclic behavior changes, including a comparison of two Rorschach records administered at a five - day interval on a "bad" and a "good" day. The comparison covers productivity, utilization of original intellectual endowment, ego strength, affective life, social adaptation, aggression, feelings of insecurity, and intellectual viscosity. The weakening of the ego's ability to control and canalize instinctual energy on the "good" day is shown to be coupled with the emergence of sexual phantasies allied with considerable aggressive tendencies. This shift on the psychodiagnostic level appears as being correlated with a shift in certain ego functions such as vision, hearing, concentration, ability to work, sexual activity, etc. on the clinical level. Loosening of control leads to hypomanic behavior, loss of inhibition up to lack of decorum, but also to a marked increase of feelings of insecurity. While a truly manic-depressive type of disturbance could be ruled out, pt. was diagnosed as a neurotic and markedly narcissistic personality showing considerable intellectual viscosity. Special emphasis was placed upon the paradoxical discrepancy between pt's subjective evaluation of his feeling of well-being and the concurring psychodiagnostic and clinical findings.

Bibliography

1. Arndt, M. Ueber taeglichen (24-stuendigen) Wechsel psychischer Krankheitszustaende. *Allg. Zschr. f. Psychiat.* 92, 1930.
2. Bohm, E. Lehrbuch der Rorschachdiagnostik. Bern: H. Huber, 1951.
3. Menninger-Lerchenthal, E. Arbeitshypothese fuer das Studium des 24-stuendigen Phasenwechsels bei Psychosen. *Allg. Zschr. f. Psychiat.*, 95, 1935, 246-256.
4. Menninger-Lerchenthal, E. Rhythmik und Periodik in der Psychopathologie. *Wien. Zschr. f. Nervenheilk. u. d. Grenzgeb.*, 1950, 3, 261-294.
5. Richter, C. P. Two-day Cycles of Alternating Good and Bad Behaviour in Psychotic Patients. *Arch. Neurol. Psychiat.*, 1938, 39, 587-598.
6. Richter, C. P. Behaviour and Metabolic Cycles in Man and their Experimental Production in Animals. In Hoch, P. R. and Zubin, J. Experimental Psychopathology. New York: Grune & Stratton, 1956.

A CORRECTION: A COMPARISON OF SOME CONCEPTS OF PSYCHOLOGICAL HEALTH

BENJAMIN MEHLMAN AND JANICE E. KAPLAN

Kent State University

Two errors, one minor and one serious, occurred in our "A Comparison of Some Concepts of Psychological Health," published in this journal in the April issue, 1958. The less serious error involved the reversal of P values in Table 2 for the first two comparisons. The more serious error involved the interpretation of our results.

Since the use of Chi Square in this instance was really a test of independence rather than of the significance of the difference between groups, we should have concluded that there was no reason to believe that our measures are not independent of one another; *i.e.*, that the MMPI is not measuring health in the same way as the Maslow S-I Inventory or the Reisman Questionnaire, nor is the Reisman Questionnaire measuring health in precisely the same way as the Maslow S-I Inventory. With a non-significant Chi Square we should have concluded that we must fail to reject the hypothesis of independence among our measures.

Since our measures did not select the same individuals as healthy we have no reason to doubt that the different authors referred to have different concepts of mental health in mind. Table 2 shows that the MMPI designation of health did not agree with our other two measures in about one-half the cases, and that the Reisman Questionnaire designation of autonomy failed to agree with the Maslow S-I designation in about one-third the cases. These discrepancies do not by themselves establish any of the three measures used as more valid than any of the other two, but it does mean that we cannot consider differences in conception of psychological health as semantic only.

Our own tendency is to see the Maslow and Reisman concepts (while somewhat divergent) as more promising than the more traditional approach to this problem. Not only do we avoid the problems that have plagued the current nosological scheme (such as the reliability and validity of the diagnoses, the social class and cultural differences in diagnoses, etc.) but we have a greater likelihood of focusing on health in an affirmative way. It makes more imperative the investigation of Maslow's hierarchy of needs in the self-actualized individual, the cultural transcendence of Reisman's autonomous individual, the ability to love of Fromm's productive individual, or the openness to experience, existential living and trust of one's own organism of Rogers' fully functioning individual. Translating these and other concepts into reasonably adequate terms has not yet been done, and will undoubtedly present enormous problems, but the effort seems worthwhile.

We are very grateful to Dr. H. P. Iker of the University of Rochester for having called the error in interpretation to our attention. We greatly regret having misled or confused our readers, particularly on a topic which we consider so important.

EDITORIAL OPINION

Clinical Judgment vs. Statistical Prediction

The last ten years have brought an increasingly critical scrutiny of the validity of clinical predictions no matter how achieved. The APA Michigan assessment studies and the wartime OSS studies were among the first to question the validity of standard measures when used as predictors of specialized performances. Although measures of ability have stood up better than personality and projective tests, an accumulating mass of evidence appears to indicate a general lack of validity of current methods for the purposes for which we are trying to use them.

Meehl[3] and Cronbach[1] were among the first to open up the issue of the relative validity of clinical judgment *vs.* statistical or actuarial prediction. Meehl, in particular, was very critical of unsophisticated clinical judgment and stressed the importance of attempting some sort of a statistical analysis of the reliability and validity of clinical judgment. Meehl's approach, while logically sound, oversimplified the issue by creating an artificial dichotomy of clinical judgment *vs.* statistical prediction in which he compared the results from relatively *naive* clinical judges with relativeiy sophisticated actuarial techniques.

In an excellent reappraisal of the issues, Holt[2] criticizes Meehl's position by pointing out that the comparison should be made not between *naive* clinical judgment and actuarial prediction but instead between the most appropriate combination of *sophisticated* clinical judgments and statistical predictions. Holt makes an operational analysis of the most sophisticated clinical judgment as involving five steps: (a) empiric study of behavior to be predicted, (b) decisions concerning what intervening variables need to be studied, (c) identifying items giving valid measures of intervening variables, (d) processing data measuring intervening variables, and (e) combining the processed data in formulae to yield definite predictions in each case. This methodological approach promises to bridge the gap between naive clinical judgment and actuarial prediction with the most sophisticated utilization of judgment, data and statistical analysis.

At the same time Hunt,[3] in a paper delivered at the Louisiana State University Symposium on Objective Approaches to Personality Assessment, also criticized Meehl and reported a research program demonstrating the reliability of clinical judgment when well trained, experienced clinicians are used in a carefully defined judgmental situation. His program calls for an analytic investigation of the range and limits of this demonstrated ability. While more limited than Holt's broad proposals it is thoroughly in harmony with them. Hunt does make the added point that clinical judgment is merely one example of the judgmental process and that application to the clinical judgment of our knowledge of judgment in general as gathered from the pertinent findings in psychophysics, etc., will help in refining and improving clinical prediction.

Ultimately, the validity of *either* clinical judgment *or* actuarial prediction can be no better than the validity of the theoretical discrimination concerning what dimensions it is important to measure. It appears that the failure of many of the current methods to make *valid and significant* discriminations is a function of the fact that they simply do not measure the factors which actually organize important types of personality integration. Behavior is very complexly organized, and until we have a more valid multivariate analysis of what factors actually contribute etiologically, we can never be certain that the methods used actually are measuring the desired factors. At a time when we still have no comprehensive idea of the neurophysiological foundations of mental life, and know little about brain functioning, further progress must depend upon the discrimination of what it is important to measure. Unfortunately the current neglect of descriptive psychiatry adds to our difficulties.

Instead of devising ingenious tests, such as ink blots, and then experimenting more or less blindly to discover what they actually measure according to some preconceived theory, it would appear more logical to start out to differentiate diagnostically all the possible levels at which personality integration is organized by various classes of factors. After the dimensions of psychological space have been differentiated, then it becomes possible to systematically devise tests to sample all levels of functioning. Instead of being "test-centered" as we now are, we would become concerned primarily with studying how integration is achieved and maintained in each person. Tests then would be used according to their indications and contraindications instead of using a shot-gun blast and hoping to hit the mark. In our own personal practice, we have long been using the eclectic, multi-level approach to problems of diagnosis and treatment.

In the search for ultimate validity in relating measures to criteria, we should be concerned primarily with clinical judgments which may be presumed to have some degree of validity, either on the basis of being made by experts or because predictions based on them turn out to be correct. Every clinician would admit that many of his judgments represent "educated guesses" or "hunches," while only a few represent judgments of which he is absolutely certain. If we recognize (a) that the judgments of some clinicians have higher validity than the judgments of others, and (b) that not all the judgments of one clinician have the same validity, then it is illogical to sum the judgments of either all clinicians or all the judgments of one clinician and expect to come up with any very great validity. Perhaps this is why Eysenck found that psychotherapy did not seem to be accomplishing much, *i.e.*, summing the results of 50 good therapists with those of 50 poor therapists might yield a value indicating little better than chance results. What we need is to study the clinical judgmental processes of the best clinicians and more specifically, of single judgments of proven validity, in order to discover how they were arrived at. Let us recognize that many diagnosticians and therapists are such by professional courtesy only (on the basis of having passed through a course of study), so that any set of judgments representing the pooled averages of many clinicians may obscure the most valid single decisions which we are seeking to study.

One way to clarify some of the issues involved would be to have each clinician make a statement evaluating his own capabilities in different areas and with different tools, and also to indicate the certainty with which specific clinical judgments are made. The various diagnostic and therapeutic alternatives would be stated in terms of probabilities based on clinical experience. Clinical results would then be assessed in terms of known actuarial probabilities concerning various disease entities, *i.e.*, different sets of data would compare the results of, say, the use of suggestion with anxiety hysterics and with sociopathic character disorders, which conditions are known to have vastly different prognoses. Finally, we would have to study many clinicians to discover who were making the most valid clinical judgments, under what conditions, and using what methods. Thus, the ultimate criterion of the validity of clinical judgment would stem from comparison with the known results of the best clinicians using the best methods under the best conditions. But again, we must insist that all this must depend upon discriminating *what type* of clinical judgment it is important to make. Otherwise the best clinicians and the best methods will fail.

W. A. H.
F. C. T.

References

1. Cronbach, L. J. Assessment of individual differences. In Farnsworth, P. and McNemar, Q. (Eds.), *Ann. Rev. Psychol.*, VII. Stanford, Calif.: Annual Reviews, 1956, pp. 173-196.
2. Holt, R. R. Clinical and statistical prediction: a reformation and some new data. *J. abnorm. soc. Psychol.*, 1958, *56*, 1-12.
3. Hunt, W. A. An actuarial approach to clinical judgment. Paper delivered at the Louisiana State University Symposium on "Objective Approaches to Personality Assessment." To be published by the L. S. U. Press.
4. Meehl, P. E. *Clinical vs. statistical prediction.* Minneapolis: Univer. Minnesota Press, 1954.

BOOK NOTICES

WOLPE, JOSEPH. *Psychotherapy by Reciprocal Inhibition.* Stanford, Calif.: Stanford University Press, 1958, pp. 239. $5.00

This important book advances the theory that neurotic responses may be modified in psychotherapy and other life situations by the method of *reciprocal inhibition.* Starting with the author's conviction that behavior modification can occur only through the processes of growth, lesions and learning, the classic Freudian psychoanalytic method for treating the neuroses is rejected in favor of conditioning methods based on the work of Pavlov, Tolman, Hull, Skinner and others. The method of reciprocal inhibition consists in weakening neurotic reactions by a conditioning process in which they are systematically exposed to and replaced by more socially acceptable and adaptive stimuli and responses. Thus, assertive responses are utilized to counteract socially conditioned anxieties and phobias, carefully nursed positive sexual responses are used to inhibit sexual inhibitions, and conditioned relaxation is widely used through hypnotic techniques to reduce all types of tensions. The author gives logically convincing arguments that the principle of reciprocal inhibition underlies positive therapeutic responses in other methods of psychotherapy and in normal life situations. He illustrates the method with numerous case histories and examples of systematic desensitization based on relaxation, citing a series of 210 cases treated by himself with this method, comparing the results with other studies reported in the literature. If his results can be reproduced in cross-validation experiments, a new and potent method of brief psychotherapy will have been added to our armamentarium.

PIOTROWSKI, Z. A. *Perceptanalysis.* New York: Macmillan, 1957, pp. 505. $6.75

Dr. Piotrowski is a recognized authority on the Rorschach method which he has fundamentally reworked, expanded and systematized. *Perceptanalysis* is a system for synthesizing the analysis of the formal visual features of the percepts with the symbolic meaning of the content of the percept. The contribution of perceptanalysis is considered to be in providing new and better referents permitting a more objective analysis of their interrelations with important personality variables. Piotrowski's approach to these important matters is still in terms of deductive-inductive logic, supported by clinical impressions and the as-yet sparse experimental literature in the field. We look in vain for the attempt to integrate the findings of perceptanalysis with the rapidly growing base of core personality variables which have been demonstrated by factorial methods. At a time when methods are available to establish the *factorial* validity of personality concepts used in clinical work, it seems less than adequate to fall back upon *logical* validity and *face* validity as do so many workers with projective methods. Projective psychology will continue to be an isolated realm until its coordinates are integrated by experimental-statistical findings with the charted areas of psychological space. Piotrowski advances some imaginative new interpretive methods and conclusions which still must be recognized as speculative until such time as validating evidence is forthcoming.

PATTERSON, C. H. *Counseling the Emotionally Disturbed.* New York: Harper, 1958, pp. 458. $6.00

This book presents an elementary introduction to the nature of counseling methods for the emotionally disturbed, directed primarily towards educators and vocational rehabilitation workers. It does not discuss counseling methods in detail or with illustrative materials but contents itself with discussions of the incidence of emotional disorders, the background which counselors should have, and some of the possible goals to be achieved.

Charles C Thomas . Publisher presents a

New List for Psychologists . . .

NEW 1958 BOOKS

Dominick A. Barbara -- YOUR SPEECH REVEALS YOUR PERSONALITY. "In our culture," says Doctor Barbara, "there is a premium on words." A chief purpose of this book is to provide a keener understanding of communicative behavior. As the therapist increases his own awareness of the importance of communicative material, he then shares more fruitfully and constructively in the therapeutic situation. Pub. '58, 190 pp., $5.50

Clinton G. Brown and Rayford T. Saucer -- ELECTRONIC INSTRUMENTATION FOR THE BEHAVIORAL SCIENCES. This book is a happy event for it provides simple, uncomplicated information on electronic theory, circuit design, and practical instrumentation techniques. Relieve your staff from instrumentation troubles by ordering today. Pub. '58, 176 pp., 73 il., (Amer. Lec. Objective Psychiatry), $5.50

Emanuel F. Hammer -- THE CLINICAL APPLICATION OF PROJECTIVE DRAWINGS. An alert author integrates for you "the talented and vigorous new child" by bringing from the scattered fields of projective drawings the solid works of 14 experts. This unified, authoritative approach provides a definitive text on this more-and-more-frequently-employed-clinical-instrument-to-assess-personality. Pub. '58, 688 pp., 360 il., $13.50

Edith Kramer -- ART THERAPY IN A CHILDREN'S COMMUNITY. The author's original procedures are interpreted in this lucid account of her resultful art therapy gathered in 7 years at Wiltwyck School. If you are in trouble, if you want fresh ideas and new counsel and wish to study successful new influences, tested creative processes, and stimulating therapeutic movements, this worthy book will likely help you also to show desired results. Pub. '58, 256 pp., 34 il. (8 in color), (Amer. Lec. Psychology), $6.75

Ainslie Meares -- THE INTROVERT: A Psychiatric Study of Social Adjustment for Student and Layman. "This book makes you understand the introvert's personality" and, as you know, many see introverts but very few understand their personality. Herein is KNOW HOW information you can use and apply and can recommend to all concerned with the everyday life of this too frequently encountered type of individual. Pub. '58, 156 pp., $4.50

Luise J. Zucker -- EGO STRUCTURE IN PARANOID SCHIZOPHRENIA: A New Method of Evaluating Projective Material. Facilitates prognosis for it serves to make the more subtle distinction between ambulatory and hospitalized patients. A new and revealing approach to prediction of the course of the schizophrenic process. A "must" for psychologists. Pub. '58, 160 pp., 18 il., (Amer. Lec. Psychology), $5.50

JUST OFF PRESS

A NEW PATTERN FOR MENTAL HEALTH SERVICES IN A CHILDREN'S COURT by **Harris B. Peck, Molly Harrower, Mildred B. Beck,** and **John Horwitz.** This long-wanted report comes from the exciting research adventure of the court INTAKE PROJECT conducted at the New York City Court of Domestic Relations. Includes verbatim recordings, illustrative cases, practical recommendations. Pub. June '58, 96 pp., (Amer. Lec. Psychology), $3.25

SOCIAL PSYCHIATRY IN ACTION: A Therapeutic Community by **Harry A. Wilmer,** Naval Med. Research Inst. Significant! Practical! In sparkling prose! With refreshing and unique illustrations! The patient group: Acutely disturbed patients of a Naval Psychiatric Center. Description: Of group therapy as a technique. Frank reporting of techniques. New management methods. Therapy. One objective: To give you the "feeling" of a therapeutic community, of TRIALS, DIFFICULTIES, SUCCESSES. Pub. July '58, 400 pp., 50 il., $8.75

WATCH FOR

Ainslie Meares: SHAPES OF SANITY, Ready December 1958.

Charles F. Diehl: COMPENDIUM ON STUTTERING, Ready December 1958.

Dominick A. Barbara: ART OF LISTENING, Ready December 1958.

CHARLES C THOMAS • PUBLISHER • 301-327 East Lawrence Avenue • SPRINGFIELD • ILLINOIS

Provides a comprehensive study
of the causes and consequences of "objective anxiety"

PSYCHOLOGICAL STRESS

Psychoanalytic and Behavorial Studies of Surgical Patients

By Irving L. Janis, Yale University

Examines important new theories and concepts . . .

This book formulates and tests hypotheses that are pertinent to a general theory concerning human emotional reactions and thought processes in extreme situations of external threat and danger. It focuses upon understanding the psychological aspects of pain, illness, and surgery. The author includes new theoretical concepts and new research findings on major problems concerning: normal adjustment mechanisms, unconscious conflicts and modes of resolution, motivations underlying increased need for group affiliation, and conditions under which authoritative communications are accepted or rejected.

Provides an illustrative case study . . .

The first part of the book contains a detailed case study of the impact of important external events (surgery in this case) on a patient undergoing psychoanalytic treatment. Free association, fantasies, and affective reactions displayed during the crucial pre-operative and post-operative sessions are used to gain insight into unconscious motives, conflicts, and defenses activated by stressful circumstances.

Employs a combination of research techniques . . .

The major psychoanalytic hypotheses that emerge from this study of one patient are assessed in the light of additional evidence presented in the second half of the book. The additional evidence is derived from behavioral studies using two different approaches—systematic intensive interviews of a representative sample of surgical patients and correlational data from a questionnaire survey administered to several hundred young men who have recently undergone surgery.

1958 **439 pages** **$6.95**

EXPERIMENTAL DESIGN in PSYCHOLOGY and the MEDICAL SCIENCES

By A. E. Maxwell, *University of London*

1958 **147 pages** **$3.75**

Send for your examination copies

JOHN WILEY & SONS, Inc., 440 4th Ave., New York 16, N. Y.

THE MULTIDISCIPLINED APPROACH

The slow-learning child or the child with emotional difficulties needs the resources of an organization that has many different approaches to the problems involved. At Devereux the psychiatrist, the physician, the psychologist, and the educator pool experiences to give each boy and girl the environment best designed to meet his individual needs.

Students are assigned to one of nineteen semi-autonomous residential units in Pennsylvania, and similar groups in California. The entire professional resources of The Devereux Foundation are available for each individual, as needed. Hundreds of Devereux students have returned to their homes and schools after a period of special therapy.

Professional inquiries should be addressed to
John M Barclay, Director of Development,
Devereux Schools, Devon, Pennsylvania;
western residents address Keith A. Seaton, Registrar,
Devereux Schools in California, Santa Barbara, California.

THE DEVEREUX FOUNDATION

A nonprofit organization — *Santa Barbara, California*

Founded 1912 — *Devon, Pennsylvania*

SCHOOLS
COMMUNITIES
CAMPS
TRAINING
RESEARCH

HELENA T. DEVEREUX
Administrative Consultant
EDWARD L. FRENCH, Ph.D.
Director

Professional
Associate Directors
Charles M. Campbell, Jr., M.D.
Michael B. Dunn, Ph.D.
Fred E. Henry, S.T.D.
J. Clifford Scott, M.D.